



Analyse en Composantes Indépendantes Multidimensionnelles via des cumulants d'ordres variés

Hanany Ould-Baba

► To cite this version:

Hanany Ould-Baba. Analyse en Composantes Indépendantes Multidimensionnelles via des cumulants d'ordres variés. Mathématiques [math]. Université de Technologie de Compiègne, 2014. Français. NNT: . tel-01264163

HAL Id: tel-01264163

<https://hal.science/tel-01264163>

Submitted on 28 Jan 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE TECHNOLOGIE DE COMPIÈGNE
ÉCOLE DOCTORALE DE COMPIÈGNE

THÈSE

pour obtenir le titre de

Docteur en Sciences

de l'Université de Technologie de Compiègne
Mention : MATHÉMATIQUES APPLIQUÉES

Présentée et soutenue par
Hanany OULD-BABA

Analyse en Composantes Indépendantes Multidimensionnelles via des cumulants d'ordres variés

Thèse dirigée par Jérôme ANTONI et Vincent ROBIN
préparée entre les laboratoires L.M.A.C et ROBERVAL
soutenue le : 11/06/2014

Jury :

| | | | |
|----------------------|------------------|---|--|
| <i>Rapporteurs :</i> | Hichem SNOUSSI | - | Université de Technologie de Troyes |
| | Tõnu KOLLO | - | University of Tartu |
| <i>Directeurs :</i> | Jérôme ANTONI | - | Université de Lyon |
| | Vincent ROBIN | - | Université de Technologie de Compiègne |
| <i>Examineurs :</i> | Yannick DEVILLE | - | Université Paul Sabatier Toulouse 3 |
| | Pascal CHEVALIER | - | professeur au CNAM/ Expert THALES |
| | Nikolaos LIMNIOS | - | Université de Technologie de Compiègne |

Avant-propos

Le présent document récapitule les travaux entrepris par l'auteur dans le cadre de sa préparation du doctorat conjoint entre le Laboratoire Roberval (UMR 7337, Université de Technologie de Compiègne, Centre de Recherches de Royallieu CS 60319, 60203 Compiègne Cedex, FRANCE) et le Laboratoire de Mathématiques Appliquées de Compiègne (Equipe d'accueil 2222, Université de Technologie de Compiègne, Département Génie Informatique, Centre de Recherches de Royallieu; UTC, Rue Roger Couttolenc, CS 60319, 60203 Compiègne Cedex, FRANCE) sous la direction de M. Jérôme ANTONI (Université de Lyon) et M. Vincent ROBIN (Université de Technologie de Compiègne).

Hanany OULD-BABA

مَنْ صَبَرَ غَنِمَ وَ مَنْ تَفَكَّرَ عَلِمَ .
الحسن البصري

*Tout devrait être rendu aussi simple que possible,
mais pas plus simple.*
Albert Einstein

À mes chers parents,
à ma femme, à mon fils,

Remerciements

J'exprime tout d'abord toute ma reconnaissance à mes deux Directeurs de thèse, M. Jérôme ANTONI et M. Vincent ROBIN, pour avoir accepté de diriger mes recherches, qui m'ont initié à la séparation de sources et m'ont guidé tout au long de ce passionnant travail de recherche. Je les remercie très sincèrement de leurs aides, leurs remarques et suggestions, essentielles, et leurs nombreux encouragements. Profonds remerciements pour ses contributions révélatrices à la correction et l'amélioration de mon document.

Je souhaite aussi faire part de ma gratitude aux membres du jury. Merci à M. Tõnu KOLLO (*professor, University of Tartu*) pour m'avoir fait l'honneur d'être rapporteur de ma thèse. Je suis très touché par sa présence dans ce jury. Merci beaucoup à M. Hichem SNOUSSI (*professeur à l'Université de Technologie de Troyes*) d'avoir accepté d'être rapporteur de cette thèse. Je tiens aussi à remercier M. Yannick DEVILLE (*professeur à l'Université Paul Sabatier*), qui m'a fait l'honneur de présider ce jury. Je tiens aussi beaucoup à remercier M. Pascal CHEVALIER (*professeur au CNAM/ Expert THALES*) et M. Nikolaos LIMNOS (*professeur à l'Université de Technologie de Compiègne*) d'avoir consenti à être examinateurs dans mon jury de thèse.

J'en profite aussi pour remercier tous les membres du Laboratoire Roberval et du Laboratoire de mathématiques appliquées de Compiègne qui m'ont accueilli chaleureusement et souvent conseillé.

Je tiens tout particulièrement à remercier mon frère Elghoth TOLBA et mon beau-frère Mohamed-Lemine KABER pour son soutien et son encouragement.

Toute mon amitié va aux doctorants avec qui j'ai partagé le bureau pendant ces années et avec qui j'ai eu tant de discussions fructueuses. Je souhaite enfin remercier tous mes amis en Mauritanie, en Allemagne, en France et ailleurs pour leurs nombreux encouragements.

Analyse en composantes indépendantes multidimensionnelles via des cumulants d'ordres variés

Résumé

L'auteur s'intéresse au problème de l'analyse en composantes indépendantes multidimensionnelles (ACIM) qui est la généralisation naturelle du problème ordinaire de l'analyse en composantes indépendantes (ACI). Tout d'abord, afin de faciliter l'utilisation des cumulants des ordres supérieurs, nous présentons de nouvelles formules pour le calcul matriciel des matrices de cumulants d'un vecteur aléatoire réel à partir de ses matrices de moments. Outre les opérations matricielles usuelles, ces formules utilisent uniquement le produit de Kronecker, l'opérateur **vec** et des matrices de commutation. Nous pouvons immédiatement à partir de ces formules examiner de plus près les structures particulières des matrices de cumulants et ainsi donner des résultats sur les rangs de ces matrices qui caractérisent la dépendance entre les variables aléatoires constituant le vecteur aléatoire. L'intérêt pratique principal de nos formules matricielles réside certainement dans une évaluation des cumulants beaucoup plus aisée et rapide qu'avec la méthode usuelle basée sur une utilisation répétée des formules de Leonov et Shiryaev. Dans la deuxième partie de cette thèse, nous montrons que sous les hypothèses usuelles de l'analyse en composantes indépendantes multidimensionnelles, les matrices de cumulants contractées à un ordre statistique quelconque sont toutes bloc-diagonalisables dans la même base. Nous en déduisons des algorithmes de résolution d'ACIM par bloc-diagonalisation conjointe et comparons les résultats obtenus aux ordres 3 à 6, entre eux et avec d'autres méthodes, sur quelques signaux synthétiques. Des exemples simples ont été élaborés afin de justifier la nécessité de combiner des ordres différents pour garantir la meilleure séparation. Nous prouvons aussi que le cas le plus simple à traiter est celui de mélanges de sources qui ont différentes dimensions. Dans la dernière partie de cette thèse nous proposons une famille de méthodes qui exploitent uniquement les statistiques d'ordres supérieurs à deux. Sous certaines hypothèses supplémentaires, ces méthodes permettent après un blanchiment d'ordre quatre des observations de résoudre complètement le problème ACIM bruité en bloc diagonalisant conjointement un ensemble de matrices de cumulants issues des statistiques d'ordres supérieurs strictement à quatre. Une comparaison avec les méthodes ACIM à blanchiment d'ordre deux pour la séparation des activités électriques fœtale et maternelle (mesurées via trois électrodes placées sur l'abdomen de la mère) montre que cette nouvelle famille est mieux adaptée à cette application : elles permettent une séparation quasi parfaite de ces deux contributions.

Mots clés

Analyse en composantes indépendantes, analyse en composantes indépendantes multidimensionnelles, séparation aveugle de sources, séparation aveugle de sources multidimensionnelles, moment, cumulant, matrice de cumulants, matrice de kurtosis, matrices de commutation, rang des matrices de cumulants, moments d'un vecteur gaussien, diagonalisation conjointe, bloc diagonalisation conjointe, produit de Kronecker, produit étoile ...

Multidimensional independent component analysis with higher-order cumulant matrices

Abstract

The author deals with the problem of multidimensional independent component analysis (MICA) which is the natural generalization of the ordinary problem of independent component analysis (ICA). First, in order to facilitate the use of higher-order cumulants, we present new formulas for the cumulant matrices of a real random vector from its moment matrices. In addition to the usual matrix operations, these formulas use only the Kronecker product, the **vec** operator and some commutation matrices. These formulas lend themselves to examine more closely the specific structures of cumulant matrices and provide results on the ranks of these matrices that characterize the dependence between random variables composing the random vector. The main practical interest of our matrix formulas lies in much easier cumulant evaluation and faster computation than the conventional method based on repeated use of the Leonov and Shiryaev formulas. In the second part of this thesis, we show that under the usual assumptions of the independent multidimensional component analysis, contracted cumulant matrices at any statistical order are all block diagonalizable in the same basis. We derive an algorithm for solving MICA by block diagonalization and compare the results obtained to the orders 3-6, between them and with other methods, on several synthetic signals. Simple examples are developed to justify the need to combine different levels to ensure the best separation. We also prove that the easiest case to deal with is the case of mixtures of sources that have different dimensions. In the last part of this thesis we propose a set of methods that operate only the higher-order statistics. Under certain additional assumptions, these methods are shown to completely solve the noisy MICA problem without second-order whitening by joint block diagonalization of a cumulant matrices set coming from statistics of orders strictly higher than four. A comparison with the second-order based whitening MICA methods for the separation of fetal and maternal electrical activities (measured using three electrodes placed on the mother's abdomen) shows that this new family is better suited to this application : it allow an almost perfect separation of these two contributions.

Key words

Independent component analysis, multidimensional independent component analysis, blind source separation, multidimensional blind source separation, independent subspace analysis, moment, cumulant, higher-order statistic, cumulant matrix, kurtosis matrix, commutation matrix, joint diagonalization, joint block diagonalization, Kronecker product, star product ...

Table des matières

| | | |
|----------|--|-----------|
| 1 | Introduction générale | 7 |
| 1.1 | Introduction à l'ACI | 7 |
| 1.1.1 | ACI au fil des années | 8 |
| 1.2 | ACI Multidimensionnelles (ACIM) | 11 |
| 1.2.1 | ACIM sous différents noms | 12 |
| 1.2.2 | Quelques domaines d'applications d'ACIM | 13 |
| 1.2.3 | ACIM au fil des années | 13 |
| 1.3 | Quelques définitions et résultats de base | 16 |
| 1.3.1 | Le cas ACIM de taille fixe ou k -ACIM | 16 |
| 1.3.2 | ACIM, le cas général | 17 |
| 1.4 | Autres extensions de l'ACI | 19 |
| 1.5 | Les algorithmes de bloc diagonalisation conjointe | 19 |
| 2 | Nouvelles formules matricielles des cumulants d'un vecteur aléatoire réel | 23 |
| 2.1 | Introduction | 23 |
| 2.2 | Rappels sur les matrices de moments et de cumulants | 24 |
| 2.3 | Matrices de cumulants à partir des matrices de moments d'ordre inférieur | 27 |
| 2.3.1 | Enoncé du résultat pour les ordres 2 à 6 | 27 |
| 2.3.2 | Stratégie générale de démonstration | 28 |
| 2.3.3 | Démonstration de l'expression à l'ordre 3 | 30 |
| 2.3.4 | Démonstration de l'expression à l'ordre 4 | 30 |
| 2.3.5 | Démonstration de l'expression à l'ordre 5 | 31 |
| 2.3.6 | Démonstration de l'expression à l'ordre 6 | 32 |
| 2.4 | Deux conséquences faciles | 35 |
| 2.4.1 | Majoration du rang des matrices de cumulants | 35 |
| 2.4.2 | Matrices des moments d'ordres 4 et 6 d'un vecteur gaussien | 36 |
| 2.5 | Gain en temps de calcul pour l'estimation des cumulants | 36 |
| 2.5.1 | Estimation des matrices de moments et cumulants | 36 |
| 2.5.2 | Comparaisons numériques entre les estimateurs matriciels et les estimateurs scalaires de cumulants | 37 |
| 2.6 | Conclusion | 37 |
| 3 | Analyse en composantes indépendantes multidimensionnelles via des cumulants d'ordres variés | 39 |
| 3.1 | Introduction et motivations | 39 |
| 3.2 | Simplification du modèle ACIM | 40 |
| 3.2.1 | Sur l'unicité aux indéterminations près | 40 |
| 3.2.2 | Hypothèses | 40 |
| 3.2.3 | Le blanchiment d'ordre deux | 41 |
| 3.2.4 | Notations | 41 |
| 3.3 | Matrices de cumulants | 41 |
| 3.4 | Matrice de cumulant contractée | 44 |
| 3.5 | Formulation du problème ACIM à l'ordre r comme problème de BDC | 46 |
| 3.6 | SJADE $_r$: une famille d'algorithmes pour l'ACIM | 50 |
| 3.7 | De l'intérêt de combiner des ordres différents | 52 |

| | | |
|----------|---|-----------|
| 3.8 | Vers une détermination automatique de la taille des sources ? | 53 |
| 3.8.1 | Un exemple numérique | 57 |
| 3.8.2 | Estimation sans la condition de rang de $c_4(\mathbf{x})$ | 59 |
| 3.9 | Etude expérimentale | 60 |
| 3.9.1 | Indice de performance pour l'ACIM | 60 |
| 3.9.2 | Comparaisons dans le cadre scalaire (ACI) | 60 |
| 3.9.3 | Comparaisons dans le cadre multidimensionnel (ACIM) | 63 |
| 3.10 | Conclusion | 65 |
| 4 | ACIM Robuste, par blanchiment d'ordre supérieur à 2 | 67 |
| 4.1 | Introduction | 67 |
| 4.2 | Modèle et hypothèses pour l'ACIM en présence de bruit | 68 |
| 4.3 | Une méthode de séparation robuste au bruit additif gaussien | 68 |
| 4.3.1 | Un blanchiment d'ordre quatre | 68 |
| 4.3.2 | Estimation de la matrice orthogonale manquante | 70 |
| 4.3.3 | RSJADE _r : une famille d'algorithmes pour l'ACIM robuste | 70 |
| 4.4 | Petite simulation et discussion | 71 |
| 4.5 | Si la matrice de kurtosis n'est ni positive ni négative | 72 |
| 4.6 | Applications aux signaux électrocardiogrammes | 74 |
| 4.6.1 | Interprétation dipolaire | 75 |
| 4.6.2 | Problème de l'extraction du ECG de fœtus | 75 |
| 4.6.3 | Quelques solutions du problème de l'extraction d'activité fœtale | 75 |
| 4.6.4 | Résolution du problème via nos algorithmes | 76 |
| 4.7 | Conclusion | 80 |
| A | Les expressions scalaires des cumulants d'un vecteurs aléatoire | 83 |
| B | Produit de Kronecker et produit étoile | 85 |
| B.1 | Produit de Kronecker | 85 |
| B.2 | Matrice de commutation | 85 |
| B.3 | Le produit étoile | 86 |
| C | Sur la réduction de l'ensemble de matrices des cumulants contractées à bloc diagonaliser conjointement | 89 |
| | Bibliographie | 91 |

Table des figures

| | | |
|-----|--|----|
| 1.1 | Schéma général de la relation sources-observations-estimations | 8 |
| 1.2 | Exemple : une séparation d'un mélange de deux images (noir et blanc) par la méthode d'ACI, JADE [Cardoso 1993]. | 9 |
| 1.3 | Schéma général de la relation sources-observations-estimations dans un cadre ACIM. | 12 |
| 1.4 | Exemple : une séparation d'un mélange de deux images par l'ACIM, une image couleur (de dimension 3) et une image noir et blanc (de dimension 1), l'algorithme utilisé ici est SJADE ₆ , il s'agit d'une méthode ACIM d'ordre six qui sera présentée dans le chapitre 3. | 13 |
| 2.1 | Comparaisons entre les temps d'exécution (en seconde) d'une estimation de la matrice de cumulants d'ordre cinq à droite (resp. six à gauche). En bleu, une estimation basée sur l'expression scalaire (E.S) scalaire et en vert, une estimation basée sur l'expression matricielle (E.M), pour une longueur de signal fixé à 1000 échantillons, en fonction de la dimension p du vecteur \mathbf{x} , qui varie de 1 à 10. | 37 |
| 3.1 | ACI, première simulation. | 61 |
| 3.2 | La moyenne du rapport signal à interférence (SIR) pour la matrice estimée de \mathbf{A} [dB] (ACI, troisième simulation) | 62 |
| 3.3 | ACI, troisième simulation | 63 |
| 3.4 | Box-plot (ou boîte à moustache) de $I(\mathbf{G})$ pour les algorithmes SJADE ₃ , SJADE [Theis 2007a], SJADE ₅ , SJADE ₆ , MHICA, [Theis 2005a] et MSOBI [Theis 2005a], pour 100 réalisations. | 64 |
| 3.5 | ACIM, seconde simulation. | 65 |
| 3.6 | Les sources estimées (ACIM, seconde simulation) | 66 |
| 3.7 | Test des matrices de séparation (ACIM, seconde simulation) | 66 |
| 4.1 | ACIM, simulation, indice de performance $I(\mathbf{G})$, en fonction du RSB. | 72 |
| 4.2 | ACI, simulation (kurtosis de signes différents). | 73 |
| 4.3 | Electrodes disposées en triangle d'Einthoven. | 74 |
| 4.4 | Un signal ECG typique. | 74 |
| 4.5 | Emplacement des capteurs, pour un système à 8 canaux d'acquisition des signaux ECG selon [De Lathauwer 1995], 5 sur la région abdominale et 3 sur la cage thoracique. | 76 |
| 4.6 | les 3 mesures sélectionnés à partir de la base de données [De Moor 1997]. | 76 |
| 4.7 | Les contributions ECGM (de dimension 2) et ECGF (de dimension 1) estimées par SJADE _r , ($r \in \{4, 5, 6\}$) à blanchiment d'ordre 2. | 78 |
| 4.8 | Les contributions ECGM (de dimension 2) et ECGF (de dimension 1) estimées par la famille RSJADE _r , ($r \in \{5, 6\}$) à blanchiment d'ordre 4. | 79 |

Liste des tableaux

| | | |
|-----|---|----|
| 3.1 | <i>Performances (ACI, première simulation)</i> | 61 |
| 3.2 | <i>Performances (ACI, seconde simulation)</i> | 61 |
| 3.3 | <i>Performances (ACI, troisième simulation)</i> | 62 |
| 3.4 | <i>Performances (ACIM, seconde simulation)</i> | 64 |

Liste des notations

- \otimes - le produit de Kronecker
- $*$ - le produit étoile
- \mathbf{A} - une matrice
- \mathbf{a} - un vecteur
- a - un scalaire
- $\text{pgcd}(a, b)$ - le plus grand commun diviseur de a et b
- ℓ_i - est la i -ème ligne de \mathbf{A}
- \mathbf{a}_i - est le i -ème colonne de \mathbf{A}
- $a_{i,j}$ - l'élément de \mathbf{A} appartenant à la i -ème ligne et j -ème colonne
- \mathbf{A}' - la matrice transposée de \mathbf{A}
- $\text{rang}(\mathbf{A})$ - le rang de la matrice \mathbf{A}
- $\text{Tr}(\mathbf{A})$ - la trace de la matrice \mathbf{A}
- $\text{vec} \mathbf{A}$ - le vecteur $\begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_n \end{pmatrix}$
- $\|\mathbf{A}\|_F$ - la norme de Frobenius de \mathbf{A}
- $\hat{\mathbf{A}}$ - une estimée de la matrice \mathbf{A}
- $\otimes^k \mathbf{A}$ - la matrice $\underbrace{(\mathbf{A} \otimes \cdots \otimes \mathbf{A})}_{k \text{ fois}}$
- \mathbf{I}_n - la matrice identité de taille n
- $\mathbf{1}_{n \times n}$ - la matrice ne contenant que des 1 de taille $n \times n$
- $\mathbf{0}_{n \times n}$ - la matrice ne contenant que des 0 de taille $n \times n$
- $\mathbf{K}_{n,m}$ - la matrice de commutation de taille $nm \times nm$
- $\text{diag}(\mathbf{a})$ - la matrice diagonale $\begin{pmatrix} a_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_n \end{pmatrix}$
- $\text{bdiag}(\mathbf{A}_1, \cdots, \mathbf{A}_n)$ - la matrice bloc diagonale $\begin{pmatrix} \mathbf{A}_1 & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{A}_n \end{pmatrix}$
- \mathbf{x} - vecteur aléatoire
- $\phi_{\mathbf{x}}(\mathbf{u})$ ou $\phi(\mathbf{u})$ - la fonction caractéristique du vecteur aléatoire \mathbf{x}
- $\mathbb{E}(\mathbf{x})$ - l'espérance de \mathbf{x}
- $\text{cov}(\mathbf{x})$ - la matrice covariance de \mathbf{x}
- $\mathcal{C}_{(3)}^{\mathbf{x}}$ - la matrice de skewness de \mathbf{x}
- $\mathcal{C}_{(4)}^{\mathbf{x}}$ - la matrice de kurtosis de \mathbf{x}
- $m_k(\mathbf{x})$ - la matrice moment d'ordre k de \mathbf{x}
- $c_k(\mathbf{x})$ - la matrice cumulant d'ordre k de \mathbf{x}
- $\mathcal{Q}_{(k)}^{\mathbf{x}}(\mathbf{M})$ - la matrice de cumulant d'ordre k de \mathbf{x} contractée par la matrice \mathbf{M}

Liste des acronymes

| | |
|---------|--|
| ACI | Analyse en Composantes Indépendantes |
| ACIM | Analyse en Composantes Indépendantes Multidimensionnelles |
| ACP | Analyse en Composantes Principales |
| AMUSE | Algorithm for Multiple Unknown Signals Extraction |
| BDC | Bloc Diagonalisation Conjointe |
| BIOME | Blind Identification of Overcomplete MixturEs |
| BIRTH | Blind Identification of mixtures of sources using Redundancies in the daTa Hexacovariance matrix |
| BSS | Blind Source Separation |
| c-à-d. | c'est à dire |
| cfMICA | Characteristic-function based Multidimensional ICA |
| COM2 | COntrast Maximization 2 |
| CSA | Colored subspace analysis |
| DC | Diagonalisation Conjointe |
| DCA | Dependent Component Analysis |
| DVP | Décomposition en Valeurs Propres |
| DVS | Décomposition en Valeurs Singulières |
| ECG | ElectroCardioGramme |
| EMG | ElectroMyoGrammes |
| FOBIUM | Fourth Order Blind Identification of Under-determined Mixtures |
| FOBI | Fourth Order Blind Identification |
| FOOBI | Fourth-Order-Only Blind Identification |
| HICA | Hessian ICA |
| ICA | Independent Component Analysis |
| ICAR | Independent Component Analysis using Redundancies in the quadricovariance |
| IFSA | Independent Feature Subspace Analysis |
| i.i.d. | indépendante identiquement distribuée |
| INFOMAX | INfOrmation MAXimization |
| ISA | Independent Subspace Analysis |
| JADE | Joint Approximate Diagonalization of Eigenmatrices |
| MBSS | Multidimensional Blind Source Separation |
| mdSOBI | multi-dimensional SOBI |
| MHICA | Multidimensional Hessian ICA |
| MICA | Multidimensional Independent Component Analysis |
| NGSA | Non-Gaussian Subspace Analysis |
| RSB | Rapport Signal à Bruit |
| RSOBI | Robust Second Order Blind Identification |
| SAS | Séparation Aveugle de Sources |
| SEONS | Second Order Non-stationary Source Separation |
| SJADE | Subspace JADE |
| RSJADE | Robust Subspace JADE |
| SSA | Stationary Subspace Analysis |
| SOBI | Second Order Blind Identification |
| SOS | Statistiques d'Ordres Supérieurs |
| STOTD | Simultaneous Third Order Tensor Diagonalization |
| TDSEP | Temporal Decorrelation SEParation |

Organisation du document

Cette thèse vise à introduire des solutions au problème de séparation de sources via l'analyse en composantes indépendantes multidimensionnelles en exploitant les statistiques d'ordres supérieurs; elle est organisée comme suit :

Le chapitre 1

Nous rappelons le problème original de la séparation aveugle de sources via l'analyse en composantes indépendantes (ACI). Nous faisons un petit tour d'horizon sur l'ensemble des techniques et méthodes proposées ces dernières années, leurs avantages et inconvénients. Nous formulons le problème d'analyse en composantes indépendantes multidimensionnelles (ACIM) comme une généralisation naturelle de l'ACI. Nous rappelons quelques définitions et concepts de base tel que l'irréductibilité et ses liens avec l'unicité de la solution. Dans une section indépendante, nous parlons des algorithmes de bloc diagonalisation conjointe (BDC) et nous exposons l'ensemble des algorithmes proposés dans ce contexte au cours des dix dernières années. Finalement, nous parlons des autres généralisations d'ACI qui ont parfois des liens très forts avec l'ACIM.

Le chapitre 2

Notre objectif dans ce chapitre est de donner de nouvelles formules matricielles pour définir les statistiques d'ordres supérieurs (SOS) d'un vecteur aléatoire réel tel que la quadricovariance (matrice de cumulants d'ordre quatre) ou encore l'hexacovariance (matrice de cumulants d'ordre six). La recherche de ces nouvelles formules a été conduite dans le but de nous faciliter dans la suite l'étude de structures particulières, telle que la décomposition en valeurs propres (DVP) ou encore pour donner des informations sur le rang. Ces formules conduiront facilement dans le chapitre qui suit à la formulation du problème ACIM comme un problème de bloc diagonalisation conjointe d'un ensemble des matrices issues de ces quantités. Par ailleurs, ces formules matricielles ont l'avantage de permettre une estimation plus rapide de tous les cumulants d'un vecteur aléatoire en comparaison aux formules scalaires habituellement utilisées.

Le chapitre 3

Ce chapitre est le coeur de cette thèse. Sur la base des outils nécessaires développés dans le chapitre précédent, nous attaquons le problème de l'ACIM. Nous montrons que sous les hypothèses usuelles de l'analyse en composantes indépendantes multidimensionnelles, les matrices de cumulants contractées à un ordre statistique quelconque sont toutes bloc diagonalisables dans la même base. Nous en déduisons des algorithmes de résolution d'ACIM par bloc-diagonalisation conjointe et comparons les résultats obtenus aux ordres trois à six, entre eux et avec d'autres méthodes existantes, sur quelques signaux synthétiques. Nous donnons aussi une condition suffisante pour garantir l'irréductibilité d'une source multidimensionnelle. Finalement, pour consolider nos propos, nous élaborons quelques exemples, afin de justifier l'avantage de combiner des ordres différents ou encore l'utilisation des statistiques d'ordres impairs pour la séparation aveugle de sources.

Le chapitre 4

Dans le dernier chapitre de cette thèse nous proposons pour la première fois une famille de méthodes baptisée RSJADE_r qui exploite uniquement les cumulants d'ordres supérieurs à deux. Sous certaines conditions, ces méthodes permettent après un blanchiment d'ordre supérieur des observations de résoudre complètement le problème ACIM bruité. La famille RSJADE_r a été comparée favorablement aux méthodes ACIM à blanchiment d'ordre deux. Finalement, nous testons ces méthodes pour séparer les activités électriques foetale et maternelle à partir d'un mélange des activités électriques mesurées sur l'abdomen de la mère.

Introduction générale

1.1 Introduction à l'ACI

Récemment, la séparation aveugle de sources (SAS) via l'analyse en composantes indépendantes (ACI) [Haykin 2000, Roberts 2001, Hyvärinen 2004a, Comon 2007, Comon 2010] a suscité un grand intérêt dans la communauté du traitement du signal. La SAS présente de nombreuses applications parmi lesquelles les télécommunications numériques, le traitement de la parole, le traitement des images, le génie biomédical, le radar, etc. Le problème de la séparation aveugle de sources peut-être reformulé comme suit : *une antenne de m capteurs reçoit, après propagation dans un environnement inconnu \mathbf{A} , des signaux émis par n sources inconnues, l'objectif étant de retrouver les n sources à partir de la seule connaissance des m signaux d'observations* (voir figure explicative 1.3, ainsi que l'exemple 1.2). Le traitement est alors aveugle ou non-supervisé, en d'autres termes, sans aucune information supplémentaire ni sur le mélange ni sur les sources. Sans ces hypothèses, la séparation apparaît comme un problème insoluble, c'est pourquoi la plupart des techniques de séparation reposent sur l'hypothèse d'indépendance mutuelle de sources. Dans le cas des sources indépendantes, la résolution du problème de séparation de sources est équivalente à l'ACI [Comon 1994]. L'ACI peut être vue comme une généralisation de l'analyse en composantes principales (ACP) qui transforme un vecteur aléatoire de sorte à assurer la décorrélation de ces variables aléatoires. Les difficultés de la séparation aveugle sont classées selon les natures des sources mélangées :

- signaux à valeurs réelles ou complexes,
- aléatoires ou déterministes stationnaires,
- cyclo-stationnaires ou non-stationnaires,
- gaussiennes,
- sub-gaussiennes,
- super-gaussiennes,
- indépendantes ou corrélés, i.i.d., ... ;

et le type de mélange :

- linéaire instantané,
- non linéaire instantané,
- linéaire convolutif,
- non linéaire convolutif ;

Autrement dit, il y a plusieurs paramètres à prendre en considération. Le mélange le plus simple est le mélange linéaire instantané où à tout instant les observations sont des combinaisons linéaires des sources. Dans le cas général le mélange est souvent convolutif, où à chaque instant les observations dépendent des sources aux instants précédents. Toutefois, il existe des techniques permettant de reformuler le modèle de mélange convolutif en un modèle de type instantané [Gorokhov 1997, Smaragdis 1998] ; c'est d'ailleurs pour cette raison que le problème le plus étudié en littérature est celui associé au mélange linéaire instantané. Nous précisons que dans tout ce document nous n'allons traiter que le modèle linéaire instantané. Un autre paramètre très important est la dimension m du vecteur d'observations par rapport à la dimension n du vecteur sources. On constate que le cas ($m \geq n$) appelé sur-déterminé est plus simple à résoudre que le cas sous-déterminé ($m < n$). Dans cette thèse nous nous plaçons aussi dans un contexte sur-déterminé ($m \geq n$). Par ailleurs, séparer aveuglement les sources provoque des indéterminations ; en effet, le problème peut accepter une infinité de solutions [Tong 1991, Bradaric 2000] et la contrainte d'indépendance n'est

cependant pas suffisante pour résoudre explicitement le problème. Plus précisément, [Comon 1994] a montré que sans condition supplémentaire, les sources ne peuvent être estimées qu'à une permutation et un facteur d'échelle près (les indéterminations d'ACI).

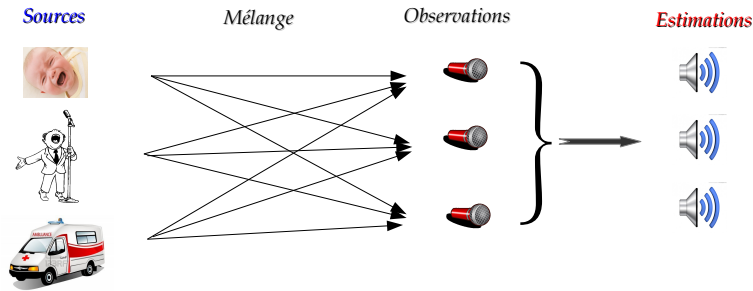


FIGURE 1.1 – Schéma général de la relation sources-observations-estimations

1.1.1 ACI au fil des années

L'ACI a tout d'abord été développée au milieu des années 80 [Hérault 1984, J. Hérault 1985] ; tout au long des années 80 le sujet est essentiellement resté un domaine de recherche plutôt français, avec une influence limitée au niveau international. Les premiers documents de l'ACI [Cardoso 1989a, Comon 1989a], ont vu le jour dans un workshop international organisé en 1989, consacré à l'analyse spectrale d'ordre supérieur (higher-order spectral analysis). En s'inspirant des travaux de C. Jutten et surtout [Jutten 1988], Cardoso [Cardoso 1989a] a introduit des méthodes algébriques exploitant la structure du tenseur de cumulants d'ordre quatre qui ont conduit au célèbre algorithme JADE (Joint Approximate Diagonalization of Eigenmatrices) [Cardoso 1993]. Plusieurs travaux pionniers d'ACI ont apparu au début des années 90, tels que [Comon 1989b, Cardoso 1989b, Jutten 1991, Pierre Comon 1991, Sorouchyari 1991]. Les travaux [Comon 1994, Tong 1991] et un peu plus tard [Bradaric 2000] ont permis de formaliser définitivement le lien entre le problème de séparation aveugle de sources et l'ACI dans le cadre du mélange linéaire sans bruit additif. Une très bonne introduction historique sur ce sujet peut être trouvée dans [Jutten 2000]. Les travaux des années 90 ont été prolongés par plusieurs auteurs, en particulier dans [Cichocki 1992, Cichocki 1996], où l'un des algorithmes d'ACI (learning algorithm ou algorithme d'apprentissage) le plus populaire à l'heure actuelle est né. L'ACI a atteint un intérêt considérable et croissant après la publication de l'approche basée sur le principe infomax [Bell 1995] dans le milieu des années 90 ; cet algorithme a été optimisé plus tard dans [Amari 1996] en utilisant la méthode du gradient naturel, augmentant ainsi la vitesse de convergence. Amari a formalisé aussi ses liens avec l'estimation du maximum de vraisemblance. Un peu plus tard, une autre version améliorée de cet algorithme est proposée dans [Lee 1999] ; cette dernière est capable de séparer un mélange contenant des sources sub-gaussiennes ou super-gaussiennes. En parallèle, Hyvärinen introduit l'algorithme du point fixe ou FastICA [Hyvärinen 1997, Hyvärinen 1999] qui est l'une des plus grandes contributions dans le domaine de la séparation de sources. Récemment, les auteurs de [Koldovský 2006] proposent une version améliorée de cet algorithme jugée asymptotiquement la plus efficace possible. L'approche matricielle JADE développée par Cardoso et Souloumiac [Cardoso 1993], basée sur les cumulants d'ordre quatre va permettre plus tard le développement des algorithmes de diagonalisation conjointes (DC) tel que [Cardoso 1996, Sorensen 2008, Pham 2000, Yeredor 2002, Yeredor 2004, Ziehe 2004, Joho 2002],

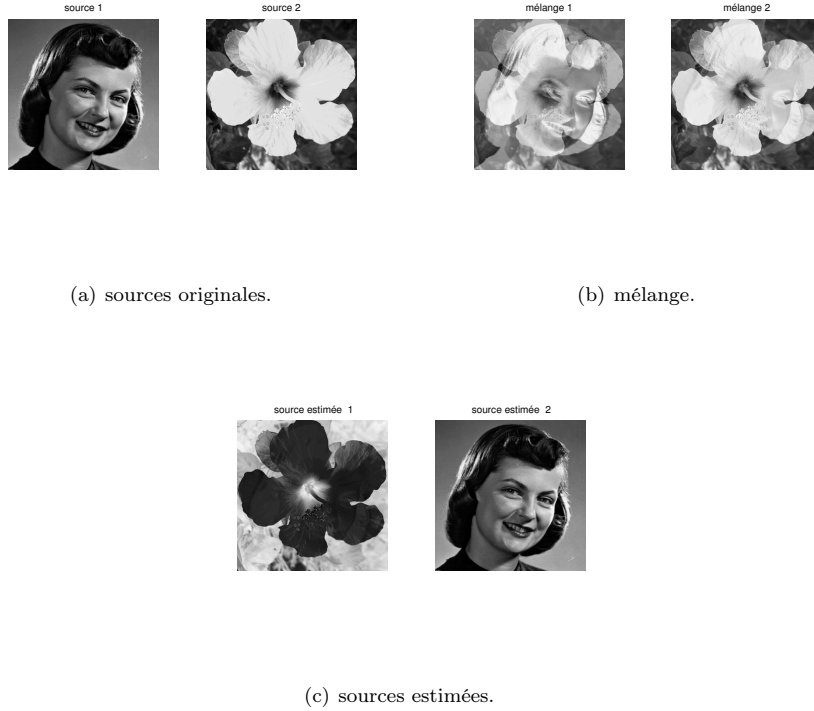


FIGURE 1.2 – Exemple : une séparation d'un mélange de deux images (noir et blanc) par la méthode d'ACI, JADE [Cardoso 1993].

donnant ainsi naissance à l'algorithme SOBI (Second Order Blind Identification) [A. Belouchrani 1997] comme extension de l'algorithme classique AMUSE (Algorithm for Multiple Unknown Signals Extraction) [Tong 1990] qui peut résoudre le problème de séparation de sources en exploitant uniquement les statistiques d'ordre deux ainsi que ses variantes : RSOBI (Robust SOBI) [Belouchrani 2000], TDSEP (Temporal Decorrelation SEPARation) [Ziehe 1998], SEONS (Second Order Non-stationary Source Separation) [Choi 2002b] et l'algorithme mdSOBI (multi-dimensional SOBI) [Theis 2004c] qui utilise des matrices d'autocovariances multidimensionnelles et mdSOBI qui a l'avantage d'utiliser les données spatiales dans toutes les directions. Dans cette famille de méthodes nous pouvons citer également les méthodes basées sur les vecteurs propres généralisés [Chang 2000, Choi 2001, Choi 2002a, Tomé 2006]. Notons bien que les algorithmes exploitant les statistiques d'ordres deux échouent complètement lorsque ils s'agit d'un mélange des sources indépendantes et identiquement distribuées, car dans ce cas les sources ne possèdent aucune diversité temporelle; c'est d'ailleurs, l'un des avantages majeurs des méthodes d'ACI d'ordres supérieures. Le premier algorithme d'ordre supérieur a été présenté par [Cardoso 1989b]; il s'agit de FOBI (Fourth Order Blind Identification). Cet algorithme a été étendu un peu plus tard par [Soon 1990]. FOBI, est probablement la méthode d'ACI d'ordre supérieur la plus simple à mettre en œuvre, en revanche, elle ne fonctionne que sous la contrainte que les kurtosis des sources sont deux à deux différents, ce qui limite les domaines d'application de cette méthode. Récemment, dans [Kollo 2008], une solution élégante à ce problème a été présentée. Mentionnons aussi, les méthodes COM1 et COM2 (CONtrast MAXimisation) [Comon 1994, Comon 2001] basées sur l'optimisation de certaines fonctions des contrastes qui exploitent

les cumulants d'ordre quatre. Un autre algorithme baptisé STOTD (Simultaneous Third Order Tensor Diagonalization) est proposé dans [De Lathauwer 1997, De Lathauwer 2001] qui suit la logique de JADE et COM2. Une fonction de contraste générale exploitant les cumulants d'ordres supérieures a été proposée dans [Moreau 2001] où il est prouvé que ce type d'approches est équivalente à la diagonalisation conjointe d'un ensemble de matrices issues des statistiques d'ordres supérieurs.

Toutes les méthodes que nous venons de rappeler ont une étape commune : l'étape de blanchiment spatial d'ordre deux ou étape de normalisation ou encore de standardisation (voir la section 3.2.3). Cette étape permet essentiellement deux choses : premièrement l'estimation du nombre de sources n , deuxièmement la transformation de la matrice de mélange en une matrice orthogonale à déterminer. Bien que ces méthodes soient très performantes dans le cadre de mélanges linéaires non bruités, elles peuvent être fortement perturbées par la présence d'un bruit gaussien de cohérence spatiale inconnue [Albera 2004], (pour plus de détail sur l'influence du blanchiment sur la séparation de sources, voir par exemple [Cardoso 1994]). Récemment, afin de surmonter cette difficulté, plusieurs algorithmes ont été conçus ; citons par exemple la méthode d'ordre deux RSOBI [Belouchrani 2000] améliorant l'algorithme SOBI [A. Belouchrani 1997] et plus précisément l'étape de blanchiment. Cette méthode a pour avantage d'être insensible asymptotiquement à un bruit temporellement décorrélié et de cohérence spatiale inconnue. En exploitant le caractère cyclostationnaire, une version modifiée de l'algorithme JADE, baptisée JADE cyclique a été proposée dans [Ferreol 1997]. Dans la même logique, les auteurs de [Antoni 2005] ont proposés un algorithme pour la séparation de sources cyclostationnaires, qui a été conçu pour être le plus robuste possible. Notons bien que le caractère cyclostationnaire a été aussi employé dans [Boustany 2005] pour traiter le cas où le nombre de sources n est inconnu (mais inférieur à m) et l'indépendance mutuelle n'est pas satisfaite, pour extraire un seul Signal d'Intérêt (SI) cyclostationnaire. L'algorithme proposé et baptisé SUBLEX exploite la cyclostationnarité de ce SI par une décomposition en sous-espaces des signaux observés via ses statistiques cycliques d'ordre deux ; l'approche développée fonctionne autant pour les mélanges instantanés que convolutifs. Cette étude a été prolongée plus tard dans [Boustany 2008] où l'algorithme RRRCR (reduced-rank cyclic regression) est proposé pour l'extraction d'un SI cyclostationnaire à partir d'un mélange potentiellement bruité. Une autre version améliorée et plus robuste de JADE baptisée Quasi-JADE est proposé récemment dans [Bonhomme 2008]. Il y a aussi la méthode ICAR [Albera 2005] qui permet sans blanchiment d'ordre deux d'identifier la matrice de mélange et qui est très robuste. En revanche, cette méthode exige la non existence d'une source gaussienne dans le mélange et aussi que les kurtosis des sources aient le même signe. Par contre, ICAR fonctionne uniquement sous la contrainte que la matrice de mélange ne contienne aucun élément nul, cette dernière contrainte a été surmontée récemment dans [Ould-Baba 2011] en élaborant un blanchiment spécial d'ordre quatre, où deux méthodes ont été présentées, la première exploitant l'ordre quatre uniquement et la deuxième combinant les statistiques d'ordres quatre et six. D'autres chercheurs se sont focalisés sur les méthodes d'ACI d'ordres supérieurs, uniquement pour résoudre le problème de séparation de sources dans un contexte sous déterminé c'est-à-dire avec plus de sources que d'observations ($n > m$). Dans cette catégorie de méthodes, nous citons d'abord la méthode classique FOABI [Cardoso 1991] qui permet sans blanchiment des observations d'identifier la matrice de mélange à partir du sous-espace signal d'ordre quatre ; cette méthode a récemment été améliorée dans [De Lathauwer 2003, De Lathauwer 2007] en se servant de la technique de diagonalisation conjointe. Il y a aussi l'extension de la méthode SOBI à l'ordre quatre, baptisée FOBIUM [Ferreol 2005], qui exploite les caractéristiques temporelles des sources afin de ne pas utiliser des statistiques d'ordres trop élevés. La méthode BIRTH [Albera 2003] exploite les cumulants d'ordre six et est une extension d'ICAR à l'ordre six. Cette méthode a été étendue à l'ordre $2q$ par l'algorithme BIOME [Albera 2004]. Notons bien que la méthode FOBIUM et la famille BIOME exploitent les redondances matricielles dans les matrices de cumulants (quadricovariance, hexacovariance, ...) des observations afin d'estimer la matrice de mélange puis les sources. Bien que ces méthodes aient été conçues pour identifier un nombre de sources supérieur au nombre de capteurs, elles sont inefficaces en pratique lorsque la matrice de mélange contient des éléments nuls ou si les cumulants des sources ont des signes différents. En revanche, cette catégorie de méthodes est par construction insensible au bruit gaussien.

En générale, les algorithmes d'ACI peuvent être divisés en deux catégories principales :

- les algorithmes basés sur les statistiques d'ordres supérieurs (SOS) et qui mesurent en quelque sorte une distance à la gaussianité ; comme d'après le théorème de la limite centrale, une somme de variables aléatoires indépendantes converge en loi vers une loi gaussienne, il en résulte que moins les variables sont gaussiennes, plus elles sont indépendantes (théorème de Granger voir [Granger 1976]) ;
- les algorithmes basés sur la théorie de l'information.

Les algorithmes exploitant les SOS se basent sur un corollaire du théorème de Darmois-Skitovitch [Darmois 1953, Skitovič 1953] qui affirme qu'au plus une source peut avoir une distribution gaussienne. En ce sens, la non-gaussianité a un rôle crucial à jouer dans l'ACI. Les mesures de non-gaussianité parmi les plus efficaces sont les cumulants d'ordres supérieurs à deux [Kendall 1977, Mac Cullagh 1987, Albera 2002, Kollo 2005]. Notons, que dans la littérature classique, la non-gaussianité est souvent mesurée par la valeur absolue du kurtosis, qui est égale à zéro si la variable aléatoire est gaussienne et supérieure à zéro pour les plupart des variables non gaussiennes. Une deuxième mesure très importante de la non-gaussianité est donnée par la néguentropie [Comon 1994, Lee 1998, Hyvarinen 1999] qui est la différence entre l'entropie d'une variable gaussienne (de même moyenne et variance que la variable aléatoire mesurée) et l'entropie de la variable mesurée. La néguentropie est toujours non négative et elle est nulle si et seulement si la distribution est gaussienne. Le problème majeur de la néguentropie est qu'elle est souvent difficile à estimer, car elle nécessiterait une estimation de la densité de probabilité. Les méthodes qui seront proposées dans cette thèse se basent uniquement sur les SOS.

1.2 ACI Multidimensionnelles (ACIM)

Dans sa présentation initiale, la séparation aveugle de sources par analyse en composantes indépendantes (ACI) suppose l'indépendance (mutuelle ou deux à deux) de sources scalaires (voir [Comon 1994]). Cette hypothèse n'est pas vérifiée pour de nombreuses applications du monde réel, ce qui peut empêcher ce modèle de récupérer certaines sources qui seraient dépendantes les unes des autres. Cette considération a naturellement amené Cardoso [Cardoso 1998] à proposer un modèle pour la séparation de sources multidimensionnelles, et non plus scalaires, en relâchant l'hypothèse assez forte d'indépendance deux à deux contre l'hypothèse suivante :

(\mathcal{H}) *Les sources inconnues peuvent être divisées en couples, triplets ou généralement en n_i -uplet, tels que les sources scalaires à l'intérieur d'un n_i -uplet donné peuvent dépendre les unes des autres, mais il n'y a pas de dépendance entre les différents n_i -uplet.*

Ainsi, Cardoso propose une généralisation du concept d'ACI ordinaire, baptisée ACIM pour ACI Multidimensionnelles (voir figure explicative 1.3, ainsi que l'exemple 1.2). Dès lors, cette généralisation a eu un intérêt croissant pour devenir l'une des approches principales du domaine du traitement aveugle des signaux. Le grand nombre de méthodes et d'algorithmes proposés ces dernières années (voir par exemple [Zoltan Szabo 2012]) témoignent de la grande importance attribuée à ce domaine de recherche. Les travaux sur l'ACIM ont porté autant d'un point de vue théorique que pratique sur la mise au point d'algorithmes de plus en plus efficaces. L'analyse en composantes indépendantes multidimensionnelles (ACIM) se formalise de la façon suivante.

$$\mathbf{x} = \mathbf{A}\mathbf{s} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \cdots & \mathbf{A}_p \end{pmatrix} \begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \vdots \\ \mathbf{s}_p \end{pmatrix} = \mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_p \quad (1.1)$$

$$\text{où : } \begin{cases} n &= \sum_{j=1}^p n_j & \text{est la dimension du vecteur source (inconnu),} \\ p &\in \mathbb{N} & \text{désigne le nombre (inconnu) de sources multidimensionnelles,} \\ \mathbf{A} &\in \mathbb{R}^{m \times n} & \text{représente la matrice de mélange (inconnue),} \\ \mathbf{s} &\in \mathbb{R}^n & \text{est le vecteur (inconnu) regroupant toutes les sources,} \\ \mathbf{s}_i &\in \mathbb{R}^{n_i} & \text{désigne la } i^{\text{ème}} \text{ source (inconnue),} \\ \mathbf{A}_i &\in \mathbb{R}^{m \times n_i} & \text{désigne la } i^{\text{ème}} \text{ matrice colonne de } \mathbf{A}, \\ \mathbf{x}_i &= \mathbf{A}_i \mathbf{s}_i \in \mathbb{R}^m & \text{désigne la } i^{\text{ème}} \text{ composante indépendante (inconnue),} \\ \mathbf{x} &\in \mathbb{R}^m & \text{représente le vecteur des observations (connu).} \end{cases} \quad (1.2)$$

On dit qu'un vecteur aléatoire \mathbf{v} est une composante indépendante (CI) d'un vecteur aléatoire \mathbf{u} , s'il existe une matrice inversible \mathbf{F} et une décomposition $\mathbf{u} = \mathbf{F} \begin{pmatrix} \mathbf{v} \\ \mathbf{w} \end{pmatrix}$ où \mathbf{v} et \mathbf{w} sont des vecteurs aléatoires indépendants. L'analyse en composantes indépendantes multidimensionnelles (ACIM) consiste en la décomposition d'un vecteur aléatoire arbitraire \mathbf{x} en CI. Aux indéterminations évidentes près (voir la sous section 3.2.1), ceci revient à trouver une estimée de la matrice de mélange \mathbf{A} (ou son inverse, dite matrice de séparation \mathbf{B}) et des sources \mathbf{s} à partir uniquement de la connaissance de \mathbf{x} . Si \mathbf{x} peut être décomposé en composantes unidimensionnelles, cela coïncide avec l'ACI ordinaire [Comon 1994]. De même, si les CI ont la même dimension k , alors cela s'appelle ACIM de taille k ou simplement k -ACIM. Evidemment, la 1-ACIM s'identifie à l'ACI [Hyvärinen 2006, Theis 2005b, Theis 2004a, Theis 2005a, Sharma 2006].

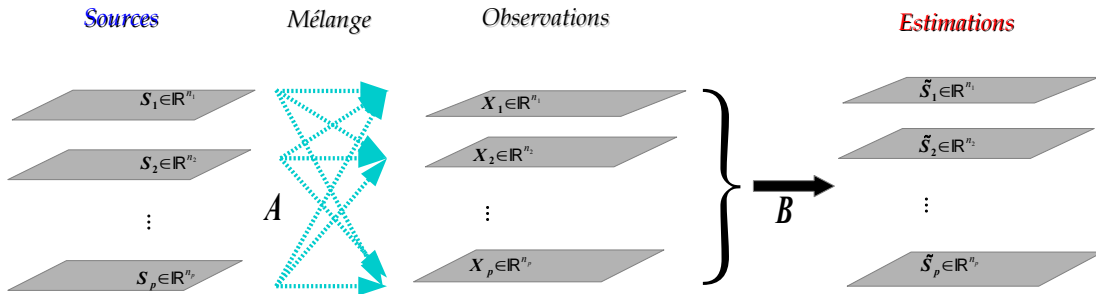


FIGURE 1.3 – Schéma général de la relation sources-observations-estimations dans un cadre ACIM.

1.2.1 ACIM sous différents noms

Signalons tout de suite que l'ACIM peut être trouvée dans la littérature sous différents noms :

- (i) independent subspace analysis (ISA) [Hyvärinen 2000],
- (ii) independent feature subspace analysis (IFSA) [Kim 2003],
- (iii) topographic ICA [Hyvärinen 2001],
- (iv) subspace ICA [Sharma 2006],
- (v) group ICA [Theis 2005a],
- (vi) independent dynamics subspace analysis [Ilin 2006],



(a) sources originales.

(b) mélange.



(c) sources estimées.

FIGURE 1.4 – Exemple : une séparation d'un mélange de deux images par l'ACIM, une image couleur (de dimension 3) et une image noir et blanc (de dimension 1), l'algorithme utilisé ici est $SJADE_6$, il s'agit d'une méthode ACIM d'ordre six qui sera présentée dans le chapitre 3.

- (vii) dependent component analysis (DCA) [Li 2010],
- (viii) tree-dependent component analysis [Bach 2003a].

1.2.2 Quelques domaines d'applications d'ACIM

Ses domaines d'applications sont très variés ; sans prétendre à l'exhaustivité, citons l'analyse d'électrocardiogrammes (ECG) [Cardoso 1998, Meinecke 2002, Theis 2005a, Sharma 2006, Chawla 2009b, Chawla 2009a, Li 2010, Yonghong 2012], le traitement de données EEG-IRMF (EEG : électro-encéphalographie, IRMF : imagerie par résonance magnétique fonctionnelle) [Ma 2010], l'analyse des images naturelles [Hyvärinen 2001], la classification des textures [Santos 2005], la reconnaissance d'action [Le 2011], la reconnaissance des visages (face recognition) [Li 2001b, Li 2001a], la segmentation de mouvement [Fan 2004], la séparation de sources (audio) mono-capteur [Casey 2000], l'analyse génétique [Kim 2003, Kim 2005, Kim 2006], l'analyse des données métaboliques [Gutch 2011] ...

1.2.3 ACIM au fil des années

Supposons que nous cherchons p composantes indépendantes multidimensionnelles, de dimension $n_i = \dim(\mathbf{s}_i)$, $i \in \llbracket 1, p \rrbracket$. La majorité des travaux proposés ont supposé afin de simplifier le problème

que les dimensions de sources multidimensionnelles sont égales [Hyvärinen 2001, Choi 2006, Bach 2003b, Theis 2005b, Theis 2004a, Theis 2005a, Hyvärinen 2006, Sharma 2006]; dans ce cas le problème est dit de taille fixe ou encore k -ACIM (avec $k = \dim(\mathbf{s}_i) = n_i$). En général, les dimensions ne sont pas connues à l'avance car elles sont inobservables directement à partir des observations. Dans cette thèse, nous nous plaçons dans un cadre général où les dimensions des sources ne sont pas (nécessairement) égales ni connues d'avance. Par conséquent, la question est de savoir comment estimer le modèle ACIM. La première approche proposée dans [Cardoso 1998] consiste à résoudre le problème en tant que problème de regroupement des éléments solutions du problème ordinaire d'ACI. Autrement dit, regrouper les éléments en fonction de leur structure de dépendance n'a un sens que si les composantes sont bien définies et peuvent être estimées avec précision; en général, nous devons exploiter les structures multidimensionnelles de sources dans la procédure d'estimation. Bien que Cardoso ait fourni dans [Cardoso 1998] le cadre général de l'étude du modèle ACIM, aucun résultat d'unicité ni de séparabilité n'a été fourni et il n'est pas clair comment appliquer l'ACIM à un vecteur aléatoire arbitraire. Un peu plus tard, une autre approche est envisagée dans [Hyvärinen 2000] où les auteurs proposent de combiner l'ACIM [Cardoso 1998] avec le principe de sous-espaces à caractéristiques invariants (Invariant feature subspaces [Kohonen 1996]). La dépendance entre les sous-espaces a été modélisée explicitement conduisant ainsi à un algorithme à direction de descente (algorithme de gradient) optimisant une fonction de vraisemblance évitant ainsi le problème classique de l'estimation de densité multidimensionnelle. Cet algorithme a été amélioré un peu plus tard en utilisant la méthode du gradient relatif dans [Choi 2006]. Bach et Jordan [Bach 2003b, Bach 2003a], introduisent un modèle ACIM qui suppose que les composants peuvent être regroupés en clusters tel qu'il y ait une dépendance entre les composants qui se retrouvent dans le même cluster et une indépendance entre les différents clusters. Ils proposent ainsi une classe d'algorithmes qui permettent de trouver ces clusters. Une technique non paramétrique basée sur l'estimation par noyau de la densité de probabilité (Kernel density estimation) a été proposée aussi dans [Samarov 2002] pour résoudre le problème d'ACI ordinaire et qui peut être étendue automatiquement (selon les auteurs) pour résoudre le problème d'ACIM. Ce problème a été aussi formulé comme un problème d'optimisation d'entropie différentielle multidimensionnelle dans les travaux présentés dans les références [Póczos 2005b, Póczos 2005a] où l'estimation de l'entropie a été effectuée en utilisant respectivement la méthode de k -plus proches voisins (en anglais k -nearest neighbors) [Póczos 2005b] et la méthode de geodesic spanning tree dans [Póczos 2005a]. Une fonction de coût pour l'ACIM combinant l'information mutuelle et l'entropie a été proposée dans [Szabó 2007], où l'information mutuelle d'un vecteur aléatoire est estimée via des différentes techniques telles que kernel generalized variance (KGV) [Bach 2003c] et kernel covariance (KC) [Gretton 2005]. L'ACIM a été aussi exploitée afin de séparer des sources corrélées comme dans [Vollgraf 2001]. La méthode FastISA [Hyvärinen 2006] est une extension multidimensionnelle de l'algorithme du point fixe FastICA [Hyvärinen 1999]. Signalons qu'une étude complémentaire de cet algorithme [Gruber 2009] a montré qu'il convergeait le plus souvent vers un minimum local et rarement vers un minimum global. Les auteurs de [Gruber 2009] ont résolu ce problème de convergence via des techniques grassmanniennes [Gruber 2006]. Récemment, une méthode ACIM qui généralise les méthodes ACI de second ordre présentées dans les références [Pham 1997, Pham 2001] et qui exploite le maximum de vraisemblance (MV) a été présentée dans [Lahat 2009]; l'approche proposée transforme le problème sous forme d'un problème de bloc diagonalisation conjointe (BDC) [Févotte 2007a, Févotte 2007b]. Par ailleurs, cette méthode est jugée de performance optimale (pour plus de détails voir [Lahat 2012a, Lahat 2012b]). D'autres ont étudié ce problème d'un point de vue géométrique en faisant intervenir la notion de variété de drapeaux (Flag Manifolds) ou encore la variété de drapeaux généralisée (Generalized Flag Manifolds); dans cette classe de méthodes nous citons les références [Nishimori 2006a, Nishimori 2006b, Nishimori 2007, Nishimori 2008]. Ces méthodes sont basées sur des optimisations de type riemannienne. Globalement, divers techniques ont été mise en oeuvre afin d'extraire les composantes indépendantes multidimensionnelles; citons : l'information mutuelle basée sur le regroupement hiérarchique [Stögbauer 2004] où les auteurs utilisent une estimation précise de l'information mutuelle (IM) basée sur un algorithme des k -plus proches voisins pour trouver les composantes les moins dépendantes dans le mélange. L'algorithme résultant permet

d'estimer les dépendances entre les composantes résiduelles des sorties ainsi que le regroupement des sorties en fonction de cette dépendance. Une fonction de contraste basée sur la mesure de dépendance (deux à deux) de Schweizer-Wolff a été optimisée dans [Kirshner 2008] pour résoudre le problème d'ACI ordinaire. Cette méthode a pu aussi fournir une solution au problème ACIM en rangeant les sorties ACI. En généralisant la notion de kurtosis d'une variable aléatoire au cas d'un vecteur aléatoire, l'auteur de [Sharma 2006] propose pour résoudre le problème ACIM un algorithme de point fixe en maximisant le kurtosis du vecteur. Toutes les méthodes que nous venons de citer jusqu'ici représentent des tentatives pour résoudre le problème d'ACIM particulier où les sources indépendantes multidimensionnelles ont une même dimension connue d'avance. L'apparition des algorithmes de bloc-diagonalisation (BDC) [Abed-Meraim 2004, Theis 2007a, Maehara 2011] d'un ensemble de matrices ont conduit à l'utilisation de statistiques d'ordre deux dans [Theis 2005a, H. Gutch 2010, Lahat 2012a] pour résoudre le problème d'ACIM en étant un problème de bloc diagonalisation d'un ensemble de matrices des corrélations retardées puis récemment comme problème de bloc-diagonalisation d'un ensemble de matrices de dispersion (scatter matrices) dans [Nordhausen 2011, Nordhausen 2012]. Toutefois, les solutions d'ACIM exploitant les statistiques d'ordre supérieur à deux, qui offrent plus d'information que les cumulants d'ordre deux, surtout lorsque certaines sources ne possèdent pas de diversité temporelle, restent rares. Les auteurs de [Castella 2007] ont étudié les cumulants d'ordres trois et quatre de sources dépendantes en considérant plusieurs cas de figures, puis ont montré que leurs cumulants peuvent avoir des propriétés intéressantes. Fabian Theis [Theis 2007a] a proposé un nouvel algorithme de bloc diagonalisation conjointe qui permet la résolution du problème de BDC lorsque les blocs n'ont pas nécessairement les mêmes dimensions, qu'il a ensuite intégré dans la méthode d'ACI ordinaire JADE [Cardoso 1993] pour engendrer une solution du problème d'ACIM. Sans formuler mathématiquement le problème d'ACIM à l'ordre 4 sous forme d'un problème de BDC, il propose de remplacer tout simplement l'étape de diagonalisation conjointe (DC) de l'ensemble des matrices de cumulants par une étape de BDC. L'algorithme résultant a été baptisé SJADE (pour Subspace-JADE). Fabian Theis, [Theis 2004a, Theis 2005a, Theis 2005b, Theis 2007a, Gutch 2007, Févotte 2007a, Gruber 2009, Maehara 2011, Gutch 2012] a proposé de nombreuses méthodes de séparation de sources multidimensionnelles via des techniques ACIM avec un intérêt marqué pour la BDC. Tout d'abord, il a commencé par étudier le problème particulier ACIM de taille fixe ou k -ACIM où toutes les sources ont la même dimension $k = \dim(\mathbf{s}_i)$ [Theis 2004a], il fournit pour la première fois un résultat de séparabilité du modèle k -ACIM, en se basant sur un théorème classique de probabilité introduit dans les années 60 dans la référence [Ghurye 1962], connue sous le nom du théorème de Darmois-Skitovitch multidimensionnel et qui est une généralisation matricielle « mais carrée » du théorème de Darmois-Skitovitch [Darmois 1953, Skitovič 1953, Skitovich 1954] (pour la version général du théorème de Darmois-Skitovitch multidimensionnel voir les références [Khatri 1972, Kagan 1973]). En se basant sur les résultats développés dans [Theis 2004a], il fournit dans [Theis 2005a], deux algorithmes pour résoudre le problème d'ACIM via la technique de bloc diagonalisation conjointe (BDC). Le premier, bloc diagonalise conjointement un ensemble de matrices de corrélations retardées, baptisé Multidimensional SOBI (MSOBI); c'est une généralisation naturelle de SOBI. Le seconde algorithme, bloc diagonalise conjointement un ensemble de matrices Hessiennes de la fonction caractéristique ou encore du logarithme de la densité de probabilité et a été baptisé Multidimensional Hessian ICA (MHICA). Les simulations informatiques ont montré que ces deux algorithmes permettent effectivement de séparer les sources indépendantes multidimensionnelles. Les travaux [Theis 2004a, Theis 2005b] ont été prolongés un peu plus tard dans [Theis 2005a] en se basant sur les études introduites dans [Eriksson 2003, Yeredor 2000, Theis 2004b]. Un peu plus tard, il avance enfin la preuve de l'unicité de la solution pour le modèle ACIM général où les sous-espaces indépendants ne sont pas nécessairement de même dimension [Theis 2007a, Gutch 2007], en remarquant tout d'abord que le modèle tel qu'il a été présenté dans [Cardoso 1998] n'est pas complet et qu'il manquait une hypothèse supplémentaire pour garantir sa séparabilité. En effet, pour n'importe quelle observation \mathbf{x} une décomposition de \mathbf{x} en sous-espaces linéairement indépendants où il existe une dépendance dans chaque sous-espaces est donnée par \mathbf{x} lui-même. Cet argument réclame naturellement une condition supplémentaire qui doit être vérifiée

par les sources afin de garantir l'unicité de la solution. Cette condition a été baptisée l'irréductibilité des sources multidimensionnelles mélangées ; nous étudierons cette notion fondamentale dans le chapitre 3 de cette thèse où nous donnons une condition suffisante pour garantir l'irréductibilité d'un vecteur aléatoire. Notons que la preuve de l'unicité de la solution ACIM a été complétée récemment dans [Gutch 2012].

1.3 Quelques définitions et résultats de base

La plupart des définitions ainsi que les résultats que nous allons donner dans cette section sont obtenus à partir des références [Khatrı 1972, Kagan 1973, Theis 2004a, Theis 2007a, Gutch 2007, Gutch 2012].

1.3.1 Le cas ACIM de taille fixe ou k -ACIM

Définition 1 ([Theis 2004a]). Une matrice partitionnée \mathbf{A} de taille $pk \times pk$ est dite k -admissible si toutes ces sous-matrices $\mathbf{A}_{i,j}$ de taille $k \times k$ sont inversibles ou nulles.

Théorème 1 ([Ghurye 1962, Theis 2004a]). Soient \mathbf{L}_1 et \mathbf{L}_2 , deux vecteurs aléatoires définies par

$$\mathbf{L}_1 = \sum_{i=1}^p \mathbf{A}_i \mathbf{s}_i \quad (1.3)$$

$$\mathbf{L}_2 = \sum_{i=1}^p \mathbf{B}_i \mathbf{s}_i \quad (1.4)$$

où les $\mathbf{s}_i \in \mathbb{R}^k$ sont des vecteurs aléatoires indépendants et les $\mathbf{A}_i \in \mathbb{R}^{k \times k}$ et $\mathbf{B}_i \in \mathbb{R}^{k \times k}$ sont des matrices inversibles. Si de plus \mathbf{L}_1 et \mathbf{L}_2 sont indépendantes, tous les \mathbf{s}_i , sont des vecteurs gaussiens.

Corollaire 2 ([Theis 2004a]). Soient \mathbf{L}_1 et \mathbf{L}_2 deux vecteurs aléatoires indépendants définis comme dans (1.3) et (1.4). Si de plus les \mathbf{A}_i et \mathbf{B}_i sont des matrices inversibles ou nulles, alors \mathbf{s}_i , tel que $\mathbf{A}_i \mathbf{B}_i \neq 0$ est gaussien.

Le corollaire 2 permet de garantir la séparabilité du problème k -ACIM si on suppose que la matrice de mélange est k -admissible. Mais avant de donner ce résultat, nous allons commencer tout d'abord par rappeler la définition d'une composante gaussienne d'un vecteur aléatoire.

Définition 2 (Composante gaussienne, [Kagan 1973]). On dit qu'un vecteur aléatoire possède une composante gaussienne si on peut l'écrire comme somme de deux vecteurs indépendants dont l'un est gaussien.

Lemme 3 ([Kagan 1973], p.310). Soit $\mathbf{x} : n \times 1$ un vecteur aléatoire. Alors on a l'alternative suivante :

- i soit \mathbf{x} est sans composante gaussienne ;
- ii soit \mathbf{x} peut s'écrire

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2, \quad (1.5)$$

où \mathbf{x}_1 et \mathbf{x}_2 sont des vecteurs aléatoires indépendants, et \mathbf{x}_2 est un vecteur gaussien de matrice de covariance maximale.

Cela veut dire que si $\mathbf{x} = \mathbf{y}_1 + \mathbf{y}_2$ est une autre décomposition de \mathbf{x} avec \mathbf{y}_1 et \mathbf{y}_2 indépendants, et \mathbf{y}_2 un vecteur gaussien, alors

$$\text{cov}[\mathbf{y}_2] - \text{cov}[\mathbf{x}_2] \text{ n'est pas définie positive.} \quad (1.6)$$

Théorème 4 (Séparabilité de k -ACIM [Theis 2004a]). Soit \mathbf{A} une matrice inversible et k -admissible et $\mathbf{s} = (\mathbf{s}'_1, \dots, \mathbf{s}'_p)' \in \mathbb{R}^{pk}$, tel que les vecteurs aléatoires $\mathbf{s}_i \in \mathbb{R}^k$ soient indépendants et sans composantes gaussiennes. Si $\mathbf{A}\mathbf{s}$ contient aussi k vecteurs aléatoires indépendants, alors $\mathbf{A} = \mathbf{P}\mathbf{D}$ où \mathbf{D} est une matrice diagonale par bloc et \mathbf{P} est une matrice de permutation.

1.3.2 ACIM, le cas général

1.3.2.1 Irréductibilité d'un vecteur aléatoire

Définition 3 ([Theis 2007a, Gutch 2007, Gutch 2012]). *Un vecteur aléatoire \mathbf{x} de dimension n est réductible si on peut l'écrire sous la forme :*

$$\mathbf{x} = \mathbf{A} \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix}, \quad (1.7)$$

où \mathbf{A} est une matrice inversible, \mathbf{y}_1 et \mathbf{y}_2 deux vecteurs aléatoires indépendants de dimensions respectives $k \neq 0$ et $n - k$. Un vecteur aléatoire qui n'est pas réductible est dit irréductible.

Remarque 5. Si $\mathbf{z} \in \mathbb{R}^n$ est un vecteur gaussien ($n > 1$), alors \mathbf{z} est réductible. Les vecteurs gaussiens sont complètement déterminés par leurs moments d'ordre un et deux, donc leur indépendance est équivalente à la décorrélation. Par ailleurs, pour tous \mathbf{z} gaussien de covariance finie, il existe une matrice inversible \mathbf{A} , tel que $\mathbf{A}\mathbf{z}$ soit décorrélée (voir [Gutch 2012]). Par conséquent, un vecteur gaussien peut toujours être réduit à une composante unidimensionnelle.

De toute évidence l'irréductibilité et la réductibilité sont conservées par les applications linéaires-inversibles. Si $\mathbf{x} = (\mathbf{x}'_1, \dots, \mathbf{x}'_p)'$, où les $\mathbf{x}_i (i = 1, \dots, p)$ sont mutuellement indépendants, cette décomposition est appelé une décomposition indépendante de \mathbf{x} . Si de plus tous les \mathbf{x}_i sont irréductibles alors dans ce cas la décomposition est appelé une décomposition irréductible.

Remarque 6. Il est facile de voir que, pour tout vecteur aléatoire $\mathbf{x} \in \mathbb{R}^n$, il existe une certaine matrice inversible $\mathbf{A} \in \mathbb{R}^{n \times n}$ telle que $\mathbf{A}\mathbf{x} = (\mathbf{x}'_1, \dots, \mathbf{x}'_p)'$ est une décomposition irréductible de \mathbf{x} . En effet, soit \mathbf{x} est lui-même irréductible, dans ce cas $\mathbf{A} = \mathbf{I}_n$, ou il existe une matrice inversible \mathbf{A} tel que $\mathbf{A}\mathbf{x} = (\mathbf{x}'_1, \mathbf{x}'_2)'$ où \mathbf{x}_1 et \mathbf{x}_2 sont indépendants. Si ces deux derniers vecteurs sont irréductibles, la condition est vérifiée, sinon on répète la même procédure pour \mathbf{x}_1 et \mathbf{x}_2 . Après un nombre fini ($p \leq \dim(\mathbf{x}) = n$) des répétitions de cette procédure nous obtiendrons la décomposition souhaitée.

Après avoir établi l'existence d'une telle décomposition, nous donnons la définition suivante.

Définition 4. Soit $\mathbf{x} \in \mathbb{R}^n$ un vecteur aléatoire et $\mathbf{A} \in \mathbb{R}^{n \times n}$ une matrice inversible, telle que $\mathbf{A}\mathbf{x} = \mathbf{s} = (\mathbf{s}'_1, \dots, \mathbf{s}'_p)'$, où

- (i) $n_i \leq n_j$ pour tout $1 \leq i \leq j \leq p$, où $n_i = \dim(\mathbf{s}_i)$,
- (ii) les vecteurs \mathbf{s}_i sont mutuellement indépendants,
- (iii) il y a au plus un \mathbf{s}_k gaussien,
- (iv) tout \mathbf{s}_k non-gaussien est irréductible,

alors, si on pose $\mathbf{n} = (n_1, \dots, n_p)$ une partition ordonnée de n , la paire (\mathbf{A}, \mathbf{n}) est appelée une analyse en sous-espaces irréductibles de \mathbf{x} (en anglais Irreducible Subspace Analysis) et les vecteurs aléatoires \mathbf{s}_i sont appelés les composantes irréductibles de (\mathbf{A}, \mathbf{n}) .

Définition 5. Dans la même logique, on dit qu'une matrice inversible \mathbf{B} est une ACIM de \mathbf{x} si $\mathbf{B}\mathbf{x}$ est une décomposition irréductible de \mathbf{x} .

1.3.2.2 Unicité de l'ACIM

Théorème 7. Soit \mathbf{x} un vecteur aléatoire dont la matrice de covariance existe et sans composante gaussienne, alors une ACIM de \mathbf{x} existe et est unique à une permutation et matrice inversible près.

La démonstration du théorème 7 a été proposée dans [Gutch 2007, Gutch 2012] en se servant des résultats suivants.

Lemme 8. Soit $\mathbf{s} = (\mathbf{s}'_1, \dots, \mathbf{s}'_p)'$ une décomposition carrée-intégrable de \mathbf{s} en composantes irréductibles et mutuellement indépendantes où chaque \mathbf{s}_i n'a aucune composante gaussienne; si $\mathbf{s} = (\mathbf{x}'_1, \mathbf{x}'_2)'$ est une autre décomposition de \mathbf{s} en composantes indépendantes, alors il existe une permutation π sur $\{1, \dots, p\}$ et un entier l , tel que

$$\mathbf{x}_1 = (\mathbf{s}'_{\pi(1)}, \dots, \mathbf{s}'_{\pi(l)})' \quad (1.8)$$

$$\mathbf{x}_2 = (\mathbf{s}'_{\pi(l+1)}, \dots, \mathbf{s}'_{\pi(p)})' \quad (1.9)$$

Lemme 9 ([Theis 2007a, Gutch 2007]). Supposons que

$$\mathbf{s} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} \quad (1.10)$$

où \mathbf{x}_1 et \mathbf{x}_2 sont deux vecteurs indépendants et les matrices \mathbf{A}_1 et \mathbf{A}_2 vérifient

$$\text{rang}(\mathbf{A}_1) + \text{rang}(\mathbf{A}_2) = \text{rang} \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 \end{pmatrix} = \dim(\mathbf{s}). \quad (1.11)$$

Alors, le vecteur aléatoire \mathbf{s} est réductible.

Lemme 10 ([Gutch 2007]). S'il existe un $k_0 \in \mathbb{N}$ tel que $\text{rang}(\mathbf{A}_{1k_0}) + \text{rang}(\mathbf{B}_{2k_0}) > \dim(\mathbf{s}_{k_0})$, alors \mathbf{s}_{k_0} a une composante gaussienne.

Le lemme 8 précédant signifie que lorsque nous avons une décomposition en composantes irréductibles et mutuellement indépendantes d'un vecteur aléatoire \mathbf{s} qui n'admet aucune composante gaussienne, n'importe quelle autre décomposition de \mathbf{s} en composantes indépendantes (pas nécessairement irréductibles) se divise le long des composantes irréductibles. Le lemme 9 permet de garantir que lorsque nous avons par exemple

$$\mathbf{x}_1 = \sum_{k=1}^p \mathbf{A}_{1k} \mathbf{s}_k \quad (1.12)$$

$$\mathbf{x}_2 = \sum_{k=1}^p \mathbf{B}_{2k} \mathbf{s}_k \quad (1.13)$$

l'une de deux matrices $\{\mathbf{A}_{1,k_0}, \mathbf{B}_{2,k_0}\}$ est forcément égale à $\mathbf{0}$; dans le cas contraire cela contredirait l'irréductibilité de la décomposition. Nous allons maintenant expliquer l'idée principale de la preuve du théorème 7 en se basant sur les trois lemmes précédants (pour la démonstration complète voir [Gutch 2012]). Il s'agit d'un raisonnement par l'absurde, en effet, si nous supposons qu'il existe un entier k_0 tel que les deux matrices \mathbf{A}_{1,k_0} et \mathbf{B}_{2,k_0} sont non nulles, c-à-d que $\text{rang}(\mathbf{A}_{1,k_0}) \neq 0$ ainsi que $\text{rang}(\mathbf{B}_{2,k_0}) \neq 0$, l'hypothèse \mathbf{A} de rang plein, entraîne que

$$\text{rang}(\mathbf{A}_{1,k_0}) + \text{rang}(\mathbf{B}_{2,k_0}) \geq \dim(\mathbf{s}_{k_0}) \quad (1.14)$$

ce qui veut dire qu'il y a deux cas à traiter

- (1) $\text{rang}(\mathbf{A}_{1,k_0}) + \text{rang}(\mathbf{B}_{2,k_0}) = \dim(\mathbf{s}_{k_0}) = n_{k_0}$
- (2) $\text{rang}(\mathbf{A}_{1,k_0}) + \text{rang}(\mathbf{B}_{2,k_0}) > \dim(\mathbf{s}_{k_0}) = n_{k_0}$

Le cas (2) est complètement résolu par le lemme 10, car nous supposons les \mathbf{s}_i sans composantes gaussiennes. Maintenant, si nous supposons que (1) vérifiée, le lemme 9 nous conduit à une contradiction avec l'hypothèse que toutes les \mathbf{s}_i sont irréductibles. Par conséquence, forcément l'une de deux matrices $\{\mathbf{A}_{1,k_0}, \mathbf{B}_{2,k_0}\}$ est égale à zéro. Finalement, le lemme 8 nous permet de conclure sur le théorème 7.

1.4 Autres extensions de l'ACI

Récemment, plusieurs travaux ont été proposés pour exclure la composante gaussienne indépendante afin d'éviter toute sorte d'indétermination additionnelle. Les auteurs de [Kawanabe 2007b, Theis 2006, Blanchard 2006, Kawanabe 2007a, Gutch 2007] ont proposé de traiter le modèle baptisé analyse en sous-espace non-gaussien (en anglais : non-gaussian subspace analysis (NGSA)). Il s'agit d'une technique qui vise à détecter un sous-espace entier non gaussien dans les données. L'analyse en composantes non-gaussiennes, se formalise de la façon suivante,

$$\mathbf{x} = \mathbf{A}\mathbf{s} = \begin{pmatrix} \mathbf{A}_n & \mathbf{A}_g \end{pmatrix} \begin{pmatrix} \mathbf{s}_n \\ \mathbf{s}_g \end{pmatrix} = \mathbf{A}_n\mathbf{s}_n + \mathbf{A}_g\mathbf{s}_g = \mathbf{x}_n + \mathbf{x}_g, \quad (1.15)$$

L'objectif de l'analyse en composantes non-gaussiennes est de trouver cette décomposition. Par ailleurs, il a été montré dans la référence [Gutch 2007] que cette décomposition est unique. Dans la même logique, d'autres généralisations ont été proposées dans la littérature ; sans être exhaustif, nous citons par exemple Colored subspace analysis (CSA) [Theis 2007b, Theis 2010], Multimodal independent component analysis [Akaho 1999], Stationary subspace analysis (SSA) [von Bünau 2009], Topographic ICA (TICA) [Hyvärinen 2001], Independent phase analysis [Almeida 2009, Almeida 2011], Variance-dependent component analysis [Hyvärinen 2004b] et Periodic component analysis [Saul 2001, Sameni 2008, Linares 2010].

1.5 Les algorithmes de bloc diagonalisation conjointe

Récemment, la communauté traitement du signal s'est beaucoup intéressée au problème de la décomposition conjointe de matrices, du fait de ses nombreuses applications en traitement d'antenne et en séparation de sources. Le premier problème à avoir focalisé l'attention est celui de la diagonalisation conjointe d'un ensemble donné de matrices par une matrice orthogonale. Cela a conduit à divers algorithmes [Cardoso 1993, Cardoso 1996, Sorensen 2008]. Ces travaux ont été prolongés afin d'être généralisés à des ensembles de tenseurs et non plus de matrices [Moreau 2001] ou encore pour surmonter la contrainte d'orthogonalité sur la matrice recherchée [Yeredor 2002, Yeredor 2004, Ziehe 2004, Joho 2002] ; une solution particulière [Pham 2000] est aussi donnée pour traiter le cas des matrices définies positives. Un second type de décomposition matricielle s'est révélé très utile en séparation aveugle de sources utilisant de représentations temps-fréquence. Il s'agit de la zéro diagonalisation conjointe [Belouchrani 2001, Chabriel 2008]. Il est également possible de voir apparaître un nouveau type de décomposition matricielle appelé bloc diagonalisation conjointe. Dans ce type de décomposition, les matrices recherchées sont maintenant bloc-diagonales ce qui s'avère particulièrement intéressant dans la résolution du problème de la séparation aveugle de mélanges convolutifs de sources [Bousbia-Salah 2001] mais également pour pouvoir estimer les directions d'arrivées des signaux [Belouchrani 1997].

Nous allons nous intéresser dans cette section au problème de bloc diagonalisation conjointe (orthogonale) d'un ensemble de matrices,

$$\mathcal{M} = \{\mathbf{M}_1, \dots, \mathbf{M}_r\} \quad (1.16)$$

où $\forall i \in \llbracket 1, r \rrbracket, \mathbf{M}_i \in \mathbb{R}^{n \times n}$. Il s'agit de la généralisation du problème ordinaire de diagonalisation conjointe (DC) [Cardoso 1993]. Par ailleurs, cette technique sera notre outil algorithmique principal pour résoudre le problème d'ACIM. Nous allons commencer par formuler le problème qui se résume à la recherche d'une matrice orthogonale \mathbf{V} qui vérifie

$$\mathbf{M}_i = \mathbf{V}\mathbf{D}_i\mathbf{V}' = \mathbf{V} \begin{pmatrix} \mathbf{D}_{i,1} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{D}_{i,p} \end{pmatrix} \mathbf{V}', \quad \forall i \in \llbracket 1, r \rrbracket. \quad (1.17)$$

Ce problème est équivalent au problème de minimisation de la fonction quadratique suivante [Févotte 2007a]

$$\mathcal{F}(\mathbf{V}, \mathbf{D}_i) = \sum_{i=1}^r \|\mathbf{M}_i - \mathbf{V}\mathbf{D}_i\mathbf{V}'\|_F \quad (1.18)$$

sous la contrainte d'orthogonalité $\mathbf{V}\mathbf{V}' = \mathbf{I}_n$, où $\|\cdot\|_F$ désigne la norme matricielle de Frobenius. Une deuxième fonction coût qui se déduit immédiatement de (1.18) peut être aussi minimisée pour résoudre le problème,

$$\mathcal{G}(\mathbf{V}, \mathbf{D}_i) = \sum_{i=1}^r \|\mathbf{V}'\mathbf{M}_i\mathbf{V} - \mathbf{D}_i\|_F \quad (1.19)$$

La matrice \mathbf{V} solution de ce problème a tout d'abord été estimée sous la contrainte que les blocs diagonaux $\mathbf{D}_{i,j}, (i, j) \in \llbracket 1, r \rrbracket \times \llbracket 1, p \rrbracket$ aient la même dimension $d = \frac{n}{p}$; plusieurs solutions ont été proposées [Févotte 2007a, Bousbia-Salah 2001, De Lathauwer 2002] en minimisant l'une de deux fonctions de coût précédentes sur l'ensemble des matrices orthogonales, paramétrées comme un produit de rotation de Givens. L'algorithme résultant est de type Jacobi, comme dans presque toutes les approches proposées, mais, les approximations faites au niveau de la fonction de coût sont différentes et il en résulte donc des algorithmes différents. Les auteurs montrent que la minimisation de la fonction coût revient à la maximisation sous la contrainte d'orthogonalité d'une forme quadratique. Une solution a été effectivement obtenue en utilisant les multiplicateurs de Lagrange, ce qui se ramène à la résolution d'un polynôme de degré quatre dans le cas de matrices réelles. Plusieurs études ont constaté des problèmes de convergence locale de ces algorithmes; ces effets ont conduit plus tard les auteurs de [Févotte 2007b] à introduire des stratégies relatives à la façon d'initialiser ces algorithmes afin d'assurer la convergence vers une bonne solution, c-à-d un minimum global. Par ailleurs, quasiment toutes les solutions proposées pour résoudre ce problème restent incapable de traiter ce problème dans un contexte général, c-à-d le cas où les blocs diagonaux n'ont pas nécessairement tous la même dimension. C'est la raison pour laquelle F.Theis a proposé dans [Theis 2007a] une solution générale au problème de BDC en se basant sur la conjecture suivante :

Conjecture 11 ([Abed-Meraïm 2004]). *Soit \mathcal{M} définie comme dans (1.16) et supposons (1.17) vrai, alors la matrice orthogonale \mathbf{E} solution du problème ordinaire de diagonalisation conjointe (DC) de l'ensemble \mathcal{M} , donnée par l'algorithme de Jacobi [Cardoso 1993], est une solution du problème de BDC à une matrice de permutation \mathbf{P} à déterminer. Autrement dit,*

$$\mathbf{V} = \mathbf{P}\mathbf{E} \quad (1.20)$$

est la solution du problème de BDC.

[Theis 2007a], a montré partiellement cette conjecture, il a proposé ensuite un algorithme de bloc diagonalisation conjointe baptisé *JBD by JD* (joint block diagonalization by joint diagonalization) qui se décompose en deux étapes :

- l'estimation de la matrice orthogonale \mathbf{E} par la méthode de Jacobi [Cardoso 1993].
- la détermination de la matrice de permutation \mathbf{P} à partir de la connaissance des produits $\mathbf{E}'\mathbf{M}_i\mathbf{E}, i \in \llbracket 1, r \rrbracket$.

Récemment, une famille de solutions a été proposée sur la notion algébrique de "matrix *-algebra" engendré par l'ensemble de matrices $\mathbf{M}_1, \mathbf{M}_2, \dots, \mathbf{M}_r$ (à bloc diagonaliser) qui avait été développée dans [Gijswijt 2010]; citons par exemple les algorithmes proposés dans les références [Murota 2010, Maehara 2010, H. Gutch 2010] ainsi que leur version améliorée baptisée *commdec* [Maehara 2011].

Il existe aussi d'autres algorithmes dont la matrice de bloc-diagonalisation n'est plus supposée orthogonale, qui ont été conçus spécialement pour éviter l'étape de blanchiment d'ordre deux, comme [Tichavsky 2012, Lahat 2012b, Nion 2011, Ghennioui 2010, Ghennioui 2007]. Par ailleurs, cette catégorie d'algorithmes nécessite la connaissance à priori de la taille de chaque bloc.

Nouvelles formules matricielles des cumulants d'un vecteur aléatoire réel

Résumé : Nous présentons de nouvelles formules pour le calcul matriciel des matrices de cumulants d'un vecteur aléatoire réel à partir de ses matrices de moments. Outre les opérations matricielles usuelles, ces formules utilisent uniquement le produit de Kronecker, l'opérateur **vec** et des matrices de commutation. Les calculs pour les ordres 5 et 6 sont détaillés ; la même méthode pourrait être utilisée pour le calcul des ordres supérieurs. Comme conséquences directes de nos formules, nous donnons (1) une majoration du rang des matrices de cumulants et (2) une expression de la matrice de moments d'ordre 6 d'un vecteur gaussien. L'intérêt pratique principal de nos formules matricielles réside certainement dans une évaluation des cumulants beaucoup plus rapide qu'avec la méthode usuelle basée sur une utilisation répétée des formules de Leonov et Shiryaev.

2.1 Introduction

Les cumulants sont très utiles pour caractériser certaines propriétés statistiques, telles que l'indépendance, le caractère gaussien ou la symétrie éventuelle de distributions statistiques. Ils ont d'abord été introduits par T.N. Thiele en 1899 sous le nom de halfinvariants (voir [Hald 2000] où le document original de Thiele est traduit en anglais). En particulier, les cumulants d'ordre supérieur ont permis la résolution de problèmes insolubles à l'ordre deux. Par exemple, en séparation aveugle de sources par l'analyse en composantes indépendantes (ACI, introduite par [Comon 1994]), les cumulants d'ordre quatre sont utilisés pour séparer des mélanges sur-déterminés (moins de sources que de capteurs), citons par exemple [Cardoso 1993] et [Albera 2005]. Récemment plusieurs chercheurs se sont intéressés aux cumulants d'ordre six ([Albera 2003]), ou plus généralement d'ordre $2k$ ([Albera 2004]), afin de séparer des mélanges potentiellement sous-déterminés (plus des sources que de capteurs). Il est bien connu que les cumulants scalaires peuvent être calculés à partir de l'évaluation des moments scalaires, via la formule classique de Leonov et Shiryaev (voir par exemple [Leonov 1959, Mac Cullagh 1987], les expressions jusqu'à l'ordre six sont données en annexe A). Il est souvent souhaitable de ranger les cumulants dans des matrices, carrées symétriques lorsqu'il s'agit d'un ordre pair, ce qui se réalise en utilisant, par exemple, la fonction **reshape** de Matlab (voir [Albera 2004]). Ce rangement facilite en effet l'exploitation de propriétés particulières telles que la redondance matricielle (voir [Albera 2003, Albera 2005, Albera 2004]), ou la diagonalisation ([Kollo 2008]).

L'inconvénient majeur de l'estimation des cumulants via la formule de Leonov et Shiryaev réside en des temps de calcul qui deviennent prohibitifs quand l'ordre des cumulants ou la dimension du vecteur aléatoire observé augmente. Pour surmonter cette difficulté, l'idée la plus simple est de chercher directement des expressions matricielles de matrices de cumulants. Une première réponse est donnée par la dérivée matricielle de [Neudecker 1969], qui permet effectivement de récupérer les cumulants d'ordres supérieur sous forme matricielles (voir par exemple, [Kollo 2005, pages 187-188]). Il reste malheureusement deux problèmes qui nous empêchent d'utiliser les expressions matricielles fournies par cette procédure :

- (i) les matrices de cumulants obtenues par dérivée matricielle ne sont pas carrées, sauf pour l'ordre deux (voir par exemple [Kollo 2005]) ;

- (ii) la manipulation des dérivées matricielles du logarithme de la fonction caractéristique devient extrêmement délicate dès l'ordre cinq et, à notre connaissance, aucune expression matricielle de cumulants n'a été donnée dans la littérature pour des ordres $r > 4$.

La difficulté (i) peut être surmontée à l'ordre quatre par une manipulation algébrique identique à celle qu'avaient utilisée [Jinadasa 1986] puis [Tracy 1993], afin d'obtenir l'expression matricielle de la matrice de moment d'ordre quatre d'un vecteur gaussien ; l'expression ainsi obtenue peut être trouvée dans [Kollo 2008, Loperfido 2014]. Par contre le point (ii) restait jusqu'à présent sans réponse. Dans ce chapitre, nous allons calculer les expressions matricielles des matrices de cumulants d'ordre cinq et six d'un vecteur aléatoire réel. Notons que la matrice de cumulants d'ordre six que nous allons obtenir sera bien symétrique.

Ce chapitre est organisé comme suit. La section §2.2 ne contient que des rappels sur les moments et cumulants d'un vecteur aléatoire, ainsi que la formule de Leonov et Shiryaev, que nous exposons ici comme conséquence de la formule de Faà di Bruno (suivant [Lukacs 1955] et [Hardy 2006]). Le théorème principal, donnant l'expression matricielle des cumulants jusqu'à l'ordre 6, est énoncé au début de la section §2.3. Avant de passer aux calculs des ordres 5 et 6, nous exposons la stratégie générale de ces calculs, qui permettrait aussi d'atteindre les ordres supérieurs à 6. Ces formules permettent d'étudier facilement les structures des matrices de cumulants, en particulier, dans la section §2.4, nous montrons comment nos formules permettent facilement d'obtenir une majoration du rang de la matrice des cumulants, puis comment obtenir sans peine la matrice des moments d'ordre 6 d'un vecteur gaussien. Notons tout de suite, que puisque nous retrouvons là une expression connue (elle figure par exemple dans [Magnus 1979], [Tracy 1993] et plus récemment dans [Schott 2003]), ce petit calcul valide l'expression obtenue pour les matrices de cumulants. La section §2.5 montre le gain en temps de calcul que procure l'utilisation de nos formules matricielles pour les cumulants par rapport à l'utilisation classique des formules de Leonov et Shiryaev.

2.2 Rappels sur les matrices de moments et de cumulants

Soit $\mathbf{x} = (x_1, \dots, x_p)'$ un vecteur aléatoire réel de dimension $p \times 1$. Sa fonction caractéristique ϕ est définie sur \mathbb{R}^p par

$$\phi(\mathbf{u}) = \mathbb{E}[\exp(i(\mathbf{u}, \mathbf{x}))] = \mathbb{E}\left[\prod_{j=1}^p \exp(iu_j x_j)\right] \quad (2.1)$$

Soit $n \in \mathbb{N}$. On dira que \mathbf{x} admet des moments d'ordre n si on a

$$\mathbb{E}[\|\mathbf{x}\|^n] < \infty, \quad (2.2)$$

où $\|\mathbf{x}\| = \left(\sum_{j=1}^p x_j^2\right)^{1/2}$ désigne la norme euclidienne de \mathbb{R}^p . Dans ce cas, (i) ϕ est n fois différentiable, (ii) tous les moments scalaires d'ordre $k \leq n$ de \mathbf{x} sont bien définis par la formule ci-dessous et on a

$$(\forall (j_1, \dots, j_k) \in \llbracket 1, p \rrbracket^k) \quad \text{mom}^{(k)}(x_{j_1}, \dots, x_{j_k}) \stackrel{\text{déf}}{=} \mathbb{E}[x_{j_1} \cdots x_{j_k}] = (-i)^k \frac{\partial^k \phi}{\partial u_{j_1} \cdots \partial u_{j_k}}(\mathbf{0}) \quad (2.3)$$

Il existe évidemment de nombreuses façons de ranger ces p^k moments scalaires d'ordre k dans une matrice. Il est naturel de viser un rangement le plus symétrique possible, et nous retenons celui déjà largement utilisé dans la littérature (voir par exemple [Schott 2003], [Jinadasa 1986], [Tracy 1993]).

Définition 6 (matrice de moments). Soit $\mathbf{x} = (x_1, \dots, x_p)'$ un vecteur aléatoire réel de dimension $p \times 1$. Ses matrices de moments d'ordres pairs et impairs, plus petits qu'un entier n tel que (2.2), sont définies respectivement par :

$$m_{2k}(\mathbf{x}) = \mathbb{E}[\otimes^k \mathbf{x} \mathbf{x}'] \quad (\text{matrice de taille } p^k \times p^k) \quad (2.4)$$

$$m_{2k+1}(\mathbf{x}) = \mathbb{E}[\otimes^k \mathbf{x} \mathbf{x}' \otimes \mathbf{x}] \quad (\text{matrice de taille } p^{k+1} \times p^k) \quad (2.5)$$

où \otimes est le produit de Kronecker et $\otimes^k \mathbf{xx}' = \underbrace{\mathbf{xx}' \otimes \cdots \otimes \mathbf{xx}'}_{k \text{ matrices } p \times p}$.

Remarquons aussi qu'il y a de nombreuses façons d'écrire ces matrices puisque, si \mathbf{a} et \mathbf{b} sont deux vecteurs taille $p \times 1$, on a

$$\mathbf{ab}' = \mathbf{a} \otimes \mathbf{b}' = \mathbf{b}' \otimes \mathbf{a} \quad (2.6)$$

Les moments scalaires (2.3) se retrouvent bien rangés dans ces matrices de moments. On le vérifie facilement en décomposant $\mathbf{x} = \sum_{j \in \llbracket 1, p \rrbracket} x_j \mathbf{e}_j$ sur la base canonique de \mathbb{R}^p et en exploitant la multilinéarité du produit de Kronecker et la linéarité de l'espérance. Pour les matrices de moments d'ordres pairs, il vient

$$\begin{aligned} m_{2k}(\mathbf{x}) &= \mathbb{E}[\otimes^k \mathbf{xx}'] = \mathbb{E} \left[\otimes^k \sum_{(j_1, j_2) \in \llbracket 1, p \rrbracket^2} x_{j_1} x_{j_2} \mathbf{e}_{j_1} \mathbf{e}_{j_2}' \right] \\ &= \mathbb{E} \left[\sum_{(j_1, j_2) \in \llbracket 1, p \rrbracket^2} x_{j_1} x_{j_2} \mathbf{e}_{j_1} \mathbf{e}_{j_2}' \otimes \cdots \otimes \sum_{(j_{2k-1}, j_{2k}) \in \llbracket 1, p \rrbracket^2} x_{j_{2k-1}} x_{j_{2k}} \mathbf{e}_{j_{2k-1}} \mathbf{e}_{j_{2k}}' \right] \\ &= \sum_{(j_1, j_2, \dots, j_{2k-1}, j_{2k}) \in \llbracket 1, p \rrbracket^{2k}} \mathbb{E}[x_{j_1} x_{j_2} \cdots x_{j_{2k-1}} x_{j_{2k}}] \mathbf{e}_{j_1} \mathbf{e}_{j_2}' \otimes \cdots \otimes \mathbf{e}_{j_{2k-1}} \mathbf{e}_{j_{2k}}' \\ &= \sum_{(j_1, j_2, \dots, j_{2k-1}, j_{2k}) \in \llbracket 1, p \rrbracket^{2k}} \text{mom}^{(2k)}(x_{j_1}, \dots, x_{j_{2k}}) \mathbf{e}_{j_1} \mathbf{e}_{j_2}' \otimes \cdots \otimes \mathbf{e}_{j_{2k-1}} \mathbf{e}_{j_{2k}}' \end{aligned} \quad (2.7)$$

De même, pour les matrices de moments d'ordres impairs, nous aurons

$$m_{2k+1}(\mathbf{x}) = \sum_{(j_1, j_2, \dots, j_{2k-1}, j_{2k}, j_{2k+1}) \in \llbracket 1, p \rrbracket^{2k+1}} \text{mom}^{(2k+1)}(x_{j_1}, \dots, x_{j_{2k+1}}) \mathbf{e}_{j_1} \mathbf{e}_{j_2}' \otimes \cdots \otimes \mathbf{e}_{j_{2k-1}} \mathbf{e}_{j_{2k}}' \otimes \mathbf{e}_{j_{2k+1}} \quad (2.8)$$

Par ailleurs, on a $\phi(\mathbf{0}) = 1$ et ϕ est continue sur \mathbb{R}^p , de sorte qu'elle ne s'annule pas dans un voisinage de $\mathbf{0} \in \mathbb{R}^p$. On considère alors la seconde fonction caractéristique, qui est donc bien définie dans un voisinage de $\mathbf{0}$, par la formule

$$\psi(\mathbf{u}) = \log(\phi(\mathbf{u})) \quad (2.9)$$

Toujours sous la condition (2.2), par composition, ψ est n fois différentiable sur un voisinage de $\mathbf{0}$; ce qui permet de définir les cumulants scalaires $k \leq n$ de \mathbf{x} par

$$(\forall (j_1, \dots, j_k) \in \llbracket 1, p \rrbracket^k) \quad \text{cum}^{(k)}(x_{j_1}, \dots, x_{j_k}) \stackrel{\text{déf}}{=} (-i)^k \frac{\partial^k \psi}{\partial u_{j_1} \cdots \partial u_{j_k}}(\mathbf{0}) \quad (2.10)$$

Il est alors tout-à-fait naturel de définir les matrices de cumulants de manière analogue aux matrices de moments.

Définition 7 (matrice de cumulants). Soit $\mathbf{x} = (x_1, \dots, x_p)'$ un vecteur aléatoire réel de dimension $p \times 1$. Soit $n \in \mathbb{N}$ et supposons que $\mathbb{E}[|\mathbf{x}|^n] < \infty$. Alors les matrices de cumulants d'ordre pair, $2k \leq n$, sont les matrices de taille $p^k \times p^k$ définies par

$$c_{2k}(\mathbf{x}) = \sum_{(j_1, j_2, \dots, j_{2k-1}, j_{2k}) \in \llbracket 1, p \rrbracket^{2k}} \text{cum}^{(2k)}(x_{j_1} x_{j_2} \cdots x_{j_{2k-1}} x_{j_{2k}}) \mathbf{e}_{j_1} \mathbf{e}_{j_2}' \otimes \cdots \otimes \mathbf{e}_{j_{2k-1}} \mathbf{e}_{j_{2k}}' \quad (2.11)$$

tandis que les matrices de cumulants d'ordre impair, $2k+1 \leq n$, sont les matrices de taille $p^{k+1} \times p^k$ définies par

$$c_{2k+1}(\mathbf{x}) = \sum_{(j_1, j_2, \dots, j_{2k}, j_{2k+1}) \in \llbracket 1, p \rrbracket^{2k+1}} \text{cum}^{(2k+1)}(x_{j_1} x_{j_2} \cdots x_{j_{2k}} x_{j_{2k+1}}) \mathbf{e}_{j_1} \mathbf{e}_{j_2}' \otimes \cdots \otimes \mathbf{e}_{j_{2k-1}} \mathbf{e}_{j_{2k}}' \otimes \mathbf{e}_{j_{2k+1}} \quad (2.12)$$

Il est clair que la formule (2.9) relie implicitement moments et cumulants et suggère deux méthodes pour calculer les cumulants en fonction des moments :

1. développer en série entière, jusqu'à l'ordre au plus n , l'expression

$$\log \left(1 + \sum_{k=1}^n \frac{(i)^k}{k!} \sum_{(j_1, \dots, j_k) \in \llbracket 1, p \rrbracket^k} \text{mom}^{(k)}(x_{j_1}, \dots, x_{j_k}) u_{j_1} \cdots u_{j_k} \right) \quad (2.13)$$

où la somme entre parenthèses est la série de Taylor tronquée à l'ordre n de $\phi(\mathbf{u})$;

2. utiliser la formule de Faà di Bruno qui donne directement une expression des dérivées partielles d'une composée d'applications.

La première méthode est celle adoptée par [Mac Cullagh 1987]. Les calculs deviennent rapidement pénibles quand l'ordre augmente : Mac Cullagh donne seulement les formules jusque l'ordre 6 ; quelques rares auteurs, comme [Albera 2002], ont poussé les calculs jusque l'ordre 8. La seconde méthode a été utilisée par [Lukacs 1955], seulement dans le cas des variables aléatoires (c'est-à-dire $p = 1$), il semble en effet qu'à cette époque l'équivalent en plusieurs variables de la formule de Faà di Bruno n'ait pas encore été calculée. C'est maintenant chose faite, et plusieurs auteurs ont proposé des formules générales pour calculer les cumulants en fonctions des moments. Il existe plusieurs façons d'écrire ces formules, selon la manière dont on rassemble les termes. La plus simple pour nous est celle démontrée par [Hardy 2006], que nous rappelons ci-dessous en utilisant nos notations.

$$\frac{\partial^k}{\partial u_{j_1} \cdots \partial u_{j_k}} f \circ \phi(\mathbf{u}) = \sum_{\pi \in P(\llbracket 1, k \rrbracket)} f^{(|\pi|)} \circ \phi(\mathbf{u}) \prod_{B \in \pi} \frac{\partial^{|B|}}{\prod_{l \in B} \partial u_{j_l}} \phi(\mathbf{u}) \quad (2.14)$$

où f est une fonction d'une seule variable, $P(\llbracket 1, k \rrbracket)$ désigne l'ensemble des partitions de $\llbracket 1, k \rrbracket$, $|E|$ désigne le cardinal d'un ensemble E et si, par exemple, $B = \{2; 3; 6\}$, on aura

$$\frac{\partial^{|B|}}{\prod_{l \in B} \partial u_{j_l}} = \frac{\partial^3}{\partial u_{j_2} \partial u_{j_3} \partial u_{j_6}}. \quad (2.15)$$

L'intérêt de la formule (2.14) réside dans le fait qu'elle est complètement développée : il n'y apparait donc pas les facteurs entiers résultant de la collecte de dérivées partielles égales lorsque des j_l coïncident. Pour calculer les cumulants en fonctions des moments, nous choisissons $f = \log$, $\mathbf{u} = \mathbf{0}$ et $\phi(\mathbf{0}) = 1$. En particulier,

$$f^{(n)}(1) = (-1)^{n-1} (n-1)! \quad (2.16)$$

et on notera, en cohérence avec (2.3),

$$\frac{\partial^{|B|}}{\prod_{l \in B} \partial u_{j_l}} \phi(\mathbf{0}) = i^{|B|} \text{mom}^{(|B|)}(x_{j_l} : l \in B) \quad (2.17)$$

En remarquant que pour tout $\pi \in P(\llbracket 1, k \rrbracket)$, on a $\sum_{B \in \pi} |B| = k$, on voit que les puissances de i se simplifient. Nous retrouvons finalement les formules appelées formules de Leonov et Shiryaev :

$$\text{cum}^{(k)}(x_{j_1}, \dots, x_{j_k}) = \sum_{\pi \in P(\llbracket 1, k \rrbracket)} (-1)^{|\pi|-1} (|\pi|-1)! \prod_{B \in \pi} \text{mom}^{(|B|)}(x_{j_l} : l \in B) \quad (2.18)$$

Pour calculer les moments en fonctions des cumulants, il suffit encore d'utiliser la formule (2.14) avec $f = \exp$, ψ à la place de Φ , $\mathbf{u} = \mathbf{0}$ et $\psi(\mathbf{0}) = 0$, pour obtenir

$$\text{mom}^{(k)}(x_{j_1}, \dots, x_{j_k}) = \sum_{\pi \in P(\llbracket 1, k \rrbracket)} \prod_{B \in \pi} \text{cum}^{(|B|)}(x_{j_l} : l \in B) \quad (2.19)$$

Nous utiliserons cette formule pour calculer $m_6(\mathbf{x})$ dans le cas où \mathbf{x} est un vecteur gaussien centré.

Evidemment les formules (2.18) permettent de calculer les matrices de cumulants $c_k(\mathbf{x})$ en fonction des matrices de moments $m_k(\mathbf{x})$ en procédant termes à termes, ce qui est extrêmement coûteux dès que l'ordre dépasse 4. Notre objectif dans ce chapitre est de donner des formules pour effectuer un tel calcul matriciellement. C'est l'objet de la section suivante.

2.3 Matrices de cumulants à partir des matrices de moments d'ordre inférieur

2.3.1 Énoncé du résultat pour les ordres 2 à 6

Pour énoncer le théorème principal de ce chapitre, nous utilisons des matrices de commutation, ainsi que l'opérateur **vec**. Les définitions et propriétés utilisées sont rassemblées dans le lemme ci-dessous, qui est un (petit) extrait de [Magnus 1979].

Lemme 12 (de commutation). *Soit un entier $p \geq 1$. On définit la matrice de commutation \mathbf{K}_{pp} , de taille $p^2 \times p^2$, par la formule*

$$\mathbf{K}_{pp} = \sum_{(i,j) \in \llbracket 1,p \rrbracket^2} \mathbf{e}_i \mathbf{e}_j' \otimes \mathbf{e}_j \mathbf{e}_i', \quad (2.20)$$

où $(\mathbf{e}_i)_{i \in \llbracket 1,p \rrbracket}$ désigne la base canonique de \mathbb{R}^p .

Définissons aussi ces deux autres matrices de commutation, toutes deux de taille $p^3 \times p^3$, par les formules :

$$\mathcal{P}_{3,2} = \mathbf{I}_p \otimes \mathbf{K}_{pp} \quad (2.21)$$

$$\mathcal{P}_{2,1} = \mathbf{K}_{pp} \otimes \mathbf{I}_p \quad (2.22)$$

Alors, si \mathbf{a} , \mathbf{b} et \mathbf{c} sont trois vecteurs de dimension $p \times 1$, nous avons les égalités suivantes

$$\mathbf{K}_{pp}(\mathbf{a} \otimes \mathbf{b}) = \mathbf{b} \otimes \mathbf{a} \quad (2.23)$$

$$\mathcal{P}_{3,2}(\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c}) = \mathbf{a} \otimes \mathbf{c} \otimes \mathbf{b} \quad (2.24)$$

$$\mathcal{P}_{2,1}(\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c}) = \mathbf{b} \otimes \mathbf{a} \otimes \mathbf{c} \quad (2.25)$$

Comme $\mathbf{K}_{pp}' = \mathbf{K}_{pp}$, on a aussi $\mathcal{P}_{3,2}' = \mathcal{P}_{3,2}$, $\mathcal{P}_{2,1}' = \mathcal{P}_{2,1}$, et donc :

$$(\mathbf{a}' \otimes \mathbf{b}') \mathbf{K}_{pp} = \mathbf{b}' \otimes \mathbf{a}' \quad (2.26)$$

$$(\mathbf{a}' \otimes \mathbf{b}' \otimes \mathbf{c}') \mathcal{P}_{3,2} = \mathbf{a}' \otimes \mathbf{c}' \otimes \mathbf{b}' \quad (2.27)$$

$$(\mathbf{a}' \otimes \mathbf{b}' \otimes \mathbf{c}') \mathcal{P}_{2,1} = \mathbf{b}' \otimes \mathbf{a}' \otimes \mathbf{c}' \quad (2.28)$$

Démonstration.

$$\begin{aligned} \mathbf{K}_{pp}(\mathbf{a} \otimes \mathbf{b}) &= \sum_{(i,j) \in \llbracket 1,p \rrbracket^2} \mathbf{e}_i \mathbf{e}_j' \otimes \mathbf{e}_j \mathbf{e}_i' (\mathbf{a} \otimes \mathbf{b}) \\ &= \sum_{(i,j) \in \llbracket 1,p \rrbracket^2} \underbrace{(\mathbf{e}_i \mathbf{e}_j' \mathbf{a})}_{a_j \mathbf{e}_i} \otimes \underbrace{(\mathbf{e}_j \mathbf{e}_i' \mathbf{b})}_{b_i \mathbf{e}_j} = \sum_{(i,j) \in \llbracket 1,p \rrbracket^2} b_i a_j \mathbf{e}_i \otimes \mathbf{e}_j = \mathbf{b} \otimes \mathbf{a} \end{aligned} \quad (2.29)$$

$$\mathcal{P}_{3,2}(\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c}) = (\mathbf{I}_p \otimes \mathbf{K}_{pp})(\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c}) = \mathbf{a} \otimes \mathbf{K}_{pp}(\mathbf{b} \otimes \mathbf{c}) = \mathbf{a} \otimes \mathbf{c} \otimes \mathbf{b} \quad (2.30)$$

$$\mathcal{P}_{2,1}(\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c}) = (\mathbf{K}_{pp} \otimes \mathbf{I}_p)(\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c}) = \mathbf{K}_{pp}(\mathbf{a} \otimes \mathbf{b}) \otimes \mathbf{c} = \mathbf{b} \otimes \mathbf{a} \otimes \mathbf{c} \quad (2.31)$$

□

Remarque 13. On peut montrer facilement que

$$\mathcal{P}_{3,2}\mathcal{P}_{2,1} = (\mathbf{I}_p \otimes \mathbf{K}_{pp})(\mathbf{K}_{pp} \otimes \mathbf{I}_p) = \mathbf{K}_{pp^2} \quad (2.32)$$

$$\mathcal{P}_{2,1}\mathcal{P}_{3,2} = (\mathcal{P}_{3,2}\mathcal{P}_{2,1})' = \mathbf{K}_{pp^2}' = \mathbf{K}_{p^2p} \quad (2.33)$$

Nous sommes maintenant en mesure d'énoncer notre résultat.

Théorème 14. Soit $\mathbf{x} = (x_1, \dots, x_p)'$ un vecteur aléatoire réel de dimension $p \times 1$, qu'on supposera, pour simplifier, centré ($\mathbb{E}[\mathbf{x}] = 0$). Pour garantir l'existence de tous les moments et cumulants considérés dans ce théorème, supposons que $\mathbb{E}[\|\mathbf{x}\|^n] < \infty$ pour un entier n au moins égal à 6.

Définissons ces quatre matrices qui ne dépendent que de p :

$$\mathbf{K} = \mathbf{I}_{p^2} + \mathbf{K}_{pp} \quad (2.34)$$

$$\mathbf{P} = \mathbf{I}_{p^3} + \mathcal{P}_{2,1} + \mathcal{P}_{3,2}\mathcal{P}_{2,1} \quad (2.35)$$

$$\mathbf{R} = \mathbf{I}_{p^3} + \mathcal{P}_{3,2} + \mathcal{P}_{2,1}\mathcal{P}_{3,2} \quad (2.36)$$

$$\mathbf{Q} = \mathbf{I}_{p^3} + \mathcal{P}_{3,2} + \mathcal{P}_{2,1} + \mathcal{P}_{2,1}\mathcal{P}_{3,2} + \mathcal{P}_{3,2}\mathcal{P}_{2,1} + \mathcal{P}_{2,1}\mathcal{P}_{3,2}\mathcal{P}_{2,1} \quad (2.37)$$

Alors les matrices de cumulants d'ordre 2 à 6 sont respectivement

$$c_2(\mathbf{x}) = m_2(\mathbf{x}) = \mathbb{E}[\mathbf{x}\mathbf{x}'] \quad (2.38)$$

$$c_3(\mathbf{x}) = m_3(\mathbf{x}) = \mathbb{E}[\mathbf{x}\mathbf{x}' \otimes \mathbf{x}] \quad (2.39)$$

$$c_4(\mathbf{x}) = m_4(\mathbf{x}) - \mathbf{K} \left\{ m_2(\mathbf{x}) \otimes m_2(\mathbf{x}) \right\} - \text{vec} m_2(\mathbf{x}) \text{vec}' m_2(\mathbf{x}) \quad (2.40)$$

$$c_5(\mathbf{x}) = m_5(\mathbf{x}) - \mathbf{R} \left\{ m_3(\mathbf{x}) \otimes m_2(\mathbf{x}) \right\} \mathbf{K} - \mathbf{P} \left\{ m_3(\mathbf{x})' \otimes \text{vec} m_2(\mathbf{x}) \right\} - \text{vec} m_3(\mathbf{x}) \text{vec}' m_2(\mathbf{x}) \quad (2.41)$$

$$c_6(\mathbf{x}) = m_6(\mathbf{x}) - \Gamma_{(4,2)}(\mathbf{x}) - \Gamma_{(3,3)}(\mathbf{x}) + 2\Gamma_{(2,2,2)}(\mathbf{x}) \quad (2.42)$$

où

$$\Gamma_{(4,2)}(\mathbf{x}) = \mathbf{R} \left\{ m_4(\mathbf{x}) \otimes m_2(\mathbf{x}) \right\} \mathbf{R}' + \left\{ \mathbf{M}_4(\mathbf{x}) \otimes \text{vec}' m_2(\mathbf{x}) \right\} \mathbf{P}' + \mathbf{P} \left\{ \mathbf{M}_4(\mathbf{x})' \otimes \text{vec} m_2(\mathbf{x}) \right\}$$

$$\Gamma_{(3,3)}(\mathbf{x}) = \mathbf{R} \left\{ m_3(\mathbf{x}) \otimes m_3(\mathbf{x})' \right\} \mathbf{P}' + \text{vec} m_3(\mathbf{x}) \text{vec}' m_3(\mathbf{x})$$

$$\Gamma_{(2,2,2)}(\mathbf{x}) = \mathbf{Q} \left\{ m_2(\mathbf{x}) \otimes m_2(\mathbf{x}) \otimes m_2(\mathbf{x}) \right\} + \mathbf{P} \left\{ m_2(\mathbf{x}) \otimes \text{vec} m_2(\mathbf{x}) \text{vec}' m_2(\mathbf{x}) \right\} \mathbf{P}'$$

$$\mathbf{M}_4(\mathbf{x}) = \mathbb{E}[\mathbf{x}\mathbf{x}' \otimes \mathbf{x} \otimes \mathbf{x}]$$

Remarque 15. La matrice $c_k(\mathbf{x})$ est la version dépliée (the unfolded version) du tenseur de cumulants $\mathcal{T}_k = \{\text{cum}(x_{i_1}, \dots, x_{i_k})\}$, voir par exemple [Qi 2007].

2.3.2 Stratégie générale de démonstration

Notre calcul consiste à

- (a) injecter les formules (2.18) qui expriment les cumulants en fonction des moments dans les formules (3.7) et (2.12) donnant les matrices de cumulants ;
- (b) exprimer les termes de la somme obtenue en fonction des matrices de moments.

Explicitons la première étape dans le cas des matrices de cumulants d'ordre pair ; c'est exactement pareil pour les ordres impairs.

$$\begin{aligned} c_{2k}(\mathbf{x}) &= \sum_{(j_1, \dots, j_{2k}) \in \llbracket 1, p \rrbracket^{2k}} \left(\sum_{\pi \in P(\llbracket 1, 2k \rrbracket)} (-1)^{|\pi|-1} (|\pi| - 1)! \prod_{B \in \pi} \text{mom}^{(|B|)}(x_{j_l} : l \in B) \right) \mathbf{e}_{j_1} \mathbf{e}'_{j_2} \otimes \dots \otimes \mathbf{e}_{j_{2k-1}} \mathbf{e}'_{j_{2k}} \\ &= \sum_{s \in P(2k)} (-1)^{|s|-1} (|s| - 1)! \Gamma_s(\mathbf{x}) \end{aligned} \quad (2.43)$$

où $P(2k)$ désigne l'ensemble des partitions de l'entier $2k$, $|s|$ désigne la longueur de s et

$$\Gamma_s(\mathbf{x}) = \sum_{\pi \in P_s(\llbracket 1, 2k \rrbracket)} \underbrace{\left(\sum_{(j_1, \dots, j_{2k}) \in \llbracket 1, p \rrbracket^{2k}} \prod_{B \in \pi} \text{mom}^{(|B|)}(x_{j_l} : l \in B) \mathbf{e}_{j_1} \mathbf{e}'_{j_2} \otimes \dots \otimes \mathbf{e}_{j_{2k-1}} \mathbf{e}'_{j_{2k}} \right)}_{T_\pi}, \quad (2.44)$$

où $P_s(\llbracket 1, 2k \rrbracket)$ désigne l'ensemble des partitions de $\llbracket 1, 2k \rrbracket$ de type s . Par exemple, pour $k = 2$, on aura

$$P(4) = \{(4); (3, 1); (2, 2); (2, 1, 1); (1, 1, 1, 1)\}, \quad (2.45)$$

et pour $s = (2, 2)$, on a $|s| = 2$ et

$$P_s(\llbracket 1, 4 \rrbracket) = \{\{\{1; 2\}; \{3; 4\}\}; \{\{1; 3\}; \{2; 4\}\}; \{\{1; 4\}; \{2; 3\}\}\}. \quad (2.46)$$

Pour la seconde étape, on conviendra de parcourir l'ensemble $P(2k)$ selon l'ordre lexicographique suggéré dans l'écriture de (2.45), de sorte que le premier terme de (2.43) sera toujours celui correspondant à $s = (2k)$, $|s| = 1$, $P_s(\llbracket 1, 2k \rrbracket) = \{\{\{1; \dots; 2k\}\}\}$, donc la partition $\pi = \{\{1; \dots; 2k\}\}$ à un seul bloc $B = \{1; \dots; 2k\}$, soit

$$\Gamma_{(2k)}(\mathbf{x}) = T_{\{\{1; \dots; 2k\}\}} = \sum_{(j_1, \dots, j_{2k}) \in \llbracket 1, p \rrbracket^{2k}} \text{mom}^{(2k)}(x_{j_l} : l \in B) \mathbf{e}_{j_1} \mathbf{e}'_{j_2} \otimes \dots \otimes \mathbf{e}_{j_{2k-1}} \mathbf{e}'_{j_{2k}} = m_{2k}(\mathbf{x}) \quad (2.47)$$

De plus, dans chaque produit $\prod_{B \in \pi}$ de (2.44), on parcourra les blocs B par ordre de taille décroissante, toujours comme le suggère (2.45). Ainsi, pour chaque partition s de l'entier $2k$, il existe une unique partition π de type s de l'ensemble $\llbracket 1, 2k \rrbracket$ qui préserve globalement l'ordre naturel sur $\llbracket 1, 2k \rrbracket$. On dira que cette partition est la partition canonique associée à s . Par exemple, la partition canonique de type $s = (2, 2)$ est $\pi_s = \{\{1; 2\}; \{3; 4\}\}$. Comme les blocs d'une partition canonique sont consécutifs, le terme correspondant est facile à calculer. De façon générale, si le type s est tel que chaque bloc possède au plus un indice impair de plus que les indices pairs, le terme correspondant à la permutation canonique de type s dans (2.44) est simplement

$$T_{\pi_s} = \sum_{(j_1, \dots, j_{2k}) \in \llbracket 1, p \rrbracket^{2k}} \prod_{B \in \pi_s} \text{mom}^{(|B|)}(x_{j_l} : l \in B) \mathbf{e}_{j_1} \otimes \mathbf{e}'_{j_2} \otimes \dots \otimes \mathbf{e}_{j_{2k-1}} \otimes \mathbf{e}'_{j_{2k}} = \bigotimes_{B \in \pi_s} m_{|B|}(\mathbf{x}) \quad (2.48)$$

Pour les partitions qui ne sont pas canoniques, nous devrons réarranger $\mathbf{e}_{j_1} \otimes \mathbf{e}'_{j_2} \otimes \dots \otimes \mathbf{e}_{j_{2k-1}} \otimes \mathbf{e}'_{j_{2k}}$ en une formule qui respecte les blocs de la partition π_s . Il n'est malheureusement pas possible d'exprimer cela matriciellement de façon systématique. Constatons tout de même qu'on peut toujours permuter les produits de Kronecker du type $\mathbf{e}_j \otimes \mathbf{e}'_k = \mathbf{e}'_k \otimes \mathbf{e}_j$, et on utilisera en particulier

$$\mathbf{e}_{j_1} \otimes \mathbf{e}'_{j_2} \otimes \dots \otimes \mathbf{e}_{j_{2k-1}} \otimes \mathbf{e}'_{j_{2k}} = \mathbf{e}_{j_1} \otimes \dots \otimes \mathbf{e}_{j_{2k-1}} \otimes \mathbf{e}'_{j_2} \otimes \dots \otimes \mathbf{e}'_{j_{2k}} = (\mathbf{e}_{j_1} \otimes \dots \otimes \mathbf{e}_{j_{2k-1}})(\mathbf{e}_{j_2} \otimes \dots \otimes \mathbf{e}_{j_{2k}})'. \quad (2.49)$$

A partir de l'ordre 5, on aura besoin de réarranger des produits de Kronecker de longueur 3 ; on utilisera pour cela le lemme de commutation.

De plus, si dans un bloc B , le nombre d'éléments impairs n'est égal ni au nombre des éléments pairs, ni au nombre des éléments pairs plus un, le facteur correspondant ne sera pas de type $m_{|B|}(\mathbf{x})$. Comme dans ce chapitre, on se limite à l'ordre 6, on s'en sortira par l'utilisation de \mathbf{vec} , de $\mathbf{M}_4(\mathbf{x})$ et de transposées. Remarquons enfin que sous l'hypothèse d'un vecteur aléatoire \mathbf{x} d'espérance nulle, tous les termes ayant en facteur un moment d'ordre 1 seront nuls. Ceux-ci correspondent aux partitions qui ont un bloc de taille 1.

2.3.3 Démonstration de l'expression à l'ordre 3

La seule partition de 3 sans bloc de taille 1 est (3), et $P_{(3)}(\llbracket 1, 3 \rrbracket) = \{\{\{1; 2; 3\}\}\}$, de sorte que :

$$c_3(\mathbf{x}) = \Gamma_{(3)}(\mathbf{x}) = \sum_{(j_1, j_2, j_3) \in \llbracket 1, p \rrbracket^3} \text{mom}^{(3)}(x_{j_1}, x_{j_2}, x_{j_3}) \mathbf{e}_{j_1} \mathbf{e}'_{j_2} \otimes \mathbf{e}_{j_3} = m_3(\mathbf{x}) \quad (2.50)$$

2.3.4 Démonstration de l'expression à l'ordre 4

Dans $P(4)$ (voir (2.45)), seuls (4) et (2, 2) n'utilisent pas 1, et on a :

$$\begin{aligned} P_{(4)}(\llbracket 1, 4 \rrbracket) &= \{\{\{1; 2; 3; 4\}\}\} \\ P_{(2,2)}(\llbracket 1, 4 \rrbracket) &= \{\{\{1; 2\}; \{3; 4\}\}; \{\{1; 3\}; \{2; 4\}\}; \{\{1; 4\}; \{2; 3\}\}\}. \end{aligned}$$

Afin d'éviter toute ambiguïté, nous désignerons chaque terme de $\Gamma_{(2,2)}(\mathbf{x})$ par T_π où π est la partition à laquelle il correspond dans (2.44).

$$c_4(\mathbf{x}) = m_4(\mathbf{x}) - \Gamma_{(2,2)}(\mathbf{x}) \quad (2.51)$$

$$\begin{aligned} &= m_4(\mathbf{x}) - \underbrace{m_2(\mathbf{x}) \otimes m_2(\mathbf{x})}_{T_{\{\{1;2\};\{3;4\}\}}} \\ &\quad - \underbrace{\sum_{(j_1, \dots, j_4) \in \llbracket 1, p \rrbracket^4} \text{mom}^{(2)}(x_{j_1}, x_{j_3}) \text{mom}^{(2)}(x_{j_2}, x_{j_4}) \mathbf{e}_{j_1} \mathbf{e}'_{j_2} \otimes \mathbf{e}_{j_3} \mathbf{e}'_{j_4}}_{T_{\{\{1;3\};\{2;4\}\}}} \\ &\quad - \underbrace{\sum_{(j_1, \dots, j_4) \in \llbracket 1, p \rrbracket^4} \text{mom}^{(2)}(x_{j_1}, x_{j_4}) \text{mom}^{(2)}(x_{j_2}, x_{j_3}) \mathbf{e}_{j_1} \mathbf{e}'_{j_2} \otimes \mathbf{e}_{j_3} \mathbf{e}'_{j_4}}_{T_{\{\{1;4\};\{2;3\}\}}} \end{aligned} \quad (2.52)$$

Pour ré-ordonner les sommes, on utilise

$$\mathbf{e}_{j_1} \mathbf{e}'_{j_2} \otimes \mathbf{e}_{j_3} \mathbf{e}'_{j_4} = \mathbf{e}_{j_1} \otimes \mathbf{e}'_{j_2} \otimes \mathbf{e}_{j_3} \otimes \mathbf{e}'_{j_4} = \mathbf{e}_{j_1} \otimes \mathbf{e}_{j_3} \otimes \mathbf{e}'_{j_2} \otimes \mathbf{e}'_{j_4} = (\mathbf{e}_{j_1} \otimes \mathbf{e}_{j_3}) \otimes (\mathbf{e}_{j_2} \otimes \mathbf{e}_{j_4})' \quad (2.53)$$

$$= \mathbf{K}_{pp}(\mathbf{e}_{j_3} \otimes \mathbf{e}_{j_1})(\mathbf{e}_{j_2} \otimes \mathbf{e}_{j_4})' = \mathbf{K}_{pp}(\mathbf{e}_{j_3} \mathbf{e}'_{j_2} \otimes \mathbf{e}_{j_1} \mathbf{e}'_{j_4}) \quad (2.54)$$

de sorte que

$$\begin{aligned} T_{\{\{1;3\};\{2;4\}\}} &= \left(\sum_{(j_1, j_3) \in \llbracket 1, p \rrbracket^2} \text{mom}^{(2)}(x_{j_1}, x_{j_3}) \mathbf{e}_{j_1} \otimes \mathbf{e}_{j_3} \right) \left(\sum_{(j_2, j_4) \in \llbracket 1, p \rrbracket^2} \text{mom}^{(2)}(x_{j_2}, x_{j_4}) \mathbf{e}_{j_2} \otimes \mathbf{e}_{j_4} \right)' \\ &= \mathbf{vec} m_2(\mathbf{x}) \mathbf{vec}' m_2(\mathbf{x}) \end{aligned} \quad (2.55)$$

$$\begin{aligned} T_{\{\{1;4\};\{2;3\}\}} &= \mathbf{K}_{pp} \left(\sum_{(j_2, j_3) \in \llbracket 1, p \rrbracket^2} \text{mom}^{(2)}(x_{j_2}, x_{j_3}) \mathbf{e}_{j_3} \mathbf{e}'_{j_2} \right) \otimes \left(\sum_{(j_1, j_4) \in \llbracket 1, p \rrbracket^2} \text{mom}^{(2)}(x_{j_1}, x_{j_4}) \mathbf{e}_{j_1} \mathbf{e}'_{j_4} \right) \\ &= \mathbf{K}_{pp} m_2(\mathbf{x}) \otimes m_2(\mathbf{x}) \end{aligned} \quad (2.56)$$

En regroupant les termes en $m_2(\mathbf{x}) \otimes m_2(\mathbf{x})$, il vient finalement

$$c_4(\mathbf{x}) = m_4(\mathbf{x}) - \mathbf{K} m_2(\mathbf{x}) \otimes m_2(\mathbf{x}) - \mathbf{vec} m_2(\mathbf{x}) \mathbf{vec}' m_2(\mathbf{x}) \quad (2.57)$$

2.3.5 Démonstration de l'expression à l'ordre 5

Les partitions de 5 qui n'utilisent pas 1 sont (5) et (3, 2).

$$\begin{aligned} P_{(5)}(\llbracket 1, 5 \rrbracket) &= \{\{\{1; 2; 3; 4; 5\}\}\} \\ P_{(3,2)}(\llbracket 1, 5 \rrbracket) &= \{\{\{1; 2; 3\}; \{4; 5\}\}; \{\{1; 2; 4\}; \{3; 5\}\}; \{\{1; 3; 4\}; \{2; 5\}\}; \{\{2; 3; 4\}; \{1; 5\}\}; \\ &\quad \{\{1; 2; 5\}; \{3; 4\}\}; \{\{1; 3; 5\}; \{2; 4\}\}; \{\{2; 3; 5\}; \{1; 4\}\}; \{\{1; 4; 5\}; \{2; 3\}\}; \\ &\quad \{\{2; 4; 5\}; \{1; 3\}\}; \{\{3; 4; 5\}; \{1; 2\}\}\}. \end{aligned}$$

On vérifie que $P_{(3,2)}(\llbracket 1, 5 \rrbracket)$ possède $\frac{5!}{3!2!} = 10$ éléments. Comme précédemment, nous désignerons chaque terme de $\Gamma_{(3,2)}(\mathbf{x})$ par T_π où π est la partition à laquelle il correspond dans (2.44).

$$c_5(\mathbf{x}) = m_5(\mathbf{x}) - \Gamma_{(3,2)}(\mathbf{x}), \quad (2.58)$$

où $\Gamma_{(3,2)}(\mathbf{x}) = \sum_{\pi \in P_{(3,2)}(\llbracket 1, 5 \rrbracket)} T_\pi$. On calcule ensuite chaque T_π en plaçant devant les termes d'ordre 3.

$$T_{\{\{1;2;3\};\{4;5\}\}} = m_3(\mathbf{x}) \otimes m_2(\mathbf{x}) \quad (2.59)$$

$$T_{\{\{1;2;4\};\{3;5\}\}} = m_3(\mathbf{x})' \otimes \text{vec} m_2(\mathbf{x}) \quad (2.60)$$

$$T_{\{\{1;3;4\};\{2;5\}\}} = m_3(\mathbf{x}) \otimes m_2(\mathbf{x}) \mathbf{K}_{pp} \quad (2.61)$$

$$T_{\{\{2;3;4\};\{1;5\}\}} = \mathcal{P}_{2,1} m_3(\mathbf{x})' \otimes \text{vec} m_2(\mathbf{x}) \quad (2.62)$$

$$T_{\{\{1;2;5\};\{3;4\}\}} = \mathcal{P}_{3,2} m_3(\mathbf{x}) \otimes m_2(\mathbf{x}) \quad (2.63)$$

$$T_{\{\{1;3;5\};\{2;4\}\}} = \text{vec} m_3(\mathbf{x}) \text{vec}' m_2(\mathbf{x}) \quad (2.64)$$

$$T_{\{\{2;3;5\};\{1;4\}\}} = \mathcal{P}_{2,1} \mathcal{P}_{3,2} m_3(\mathbf{x}) \otimes m_2(\mathbf{x}) \quad (2.65)$$

$$T_{\{\{1;4;5\};\{2;3\}\}} = \mathcal{P}_{3,2} m_3(\mathbf{x}) \otimes m_2(\mathbf{x}) \mathbf{K}_{pp} \quad (2.66)$$

$$T_{\{\{2;4;5\};\{1;3\}\}} = \mathcal{P}_{3,2} \mathcal{P}_{2,1} m_3(\mathbf{x})' \otimes \text{vec} m_2(\mathbf{x}) \quad (2.67)$$

$$T_{\{\{3;4;5\};\{1;2\}\}} = \mathcal{P}_{2,1} \mathcal{P}_{3,2} m_3(\mathbf{x}) \otimes m_2(\mathbf{x}) \mathbf{K}_{pp} \quad (2.68)$$

Le principe de calcul de ces T_π est en fait très simple. Détaillons par exemple le calcul de $T_{\{\{2;3;5\};\{1;4\}\}}$.

$$\begin{aligned} T_{\{\{2;3;5\};\{1;4\}\}} &= \sum_{(j_1, \dots, j_5) \in \llbracket 1, p \rrbracket^5} \text{mom}^{(3)}(x_{j_2}, x_{j_3}, x_{j_5}) \text{mom}^{(2)}(x_{j_1}, x_{j_4}) \mathbf{e}_{j_1} \otimes \mathbf{e}_{j_2}' \otimes \mathbf{e}_{j_3} \otimes \mathbf{e}_{j_4}' \otimes \mathbf{e}_{j_5} \\ &= \sum_{(j_1, \dots, j_5) \in \llbracket 1, p \rrbracket^5} \text{mom}^{(3)}(x_{j_2}, x_{j_3}, x_{j_5}) \text{mom}^{(2)}(x_{j_1}, x_{j_4}) (\mathbf{e}_{j_1} \otimes \mathbf{e}_{j_3} \otimes \mathbf{e}_{j_5}) \otimes (\mathbf{e}_{j_2}' \otimes \mathbf{e}_{j_4}') \\ &= \sum_{(j_1, \dots, j_5) \in \llbracket 1, p \rrbracket^5} \text{mom}^{(3)}(x_{j_2}, x_{j_3}, x_{j_5}) \text{mom}^{(2)}(x_{j_1}, x_{j_4}) \mathcal{P}_{2,1} \mathcal{P}_{3,2} (\mathbf{e}_{j_3} \otimes \mathbf{e}_{j_5} \otimes \mathbf{e}_{j_1}) \otimes (\mathbf{e}_{j_2}' \otimes \mathbf{e}_{j_4}') \\ &= \mathcal{P}_{2,1} \mathcal{P}_{3,2} \sum_{(j_1, \dots, j_5) \in \llbracket 1, p \rrbracket^5} \underbrace{\text{mom}^{(3)}(x_{j_2}, x_{j_3}, x_{j_5}) \text{mom}^{(2)}(x_{j_1}, x_{j_4})}_{=\text{mom}^{(3)}(x_{j_3}, x_{j_2}, x_{j_5})} (\mathbf{e}_{j_3} \otimes \mathbf{e}_{j_2}' \otimes \mathbf{e}_{j_5}) \otimes (\mathbf{e}_{j_1} \otimes \mathbf{e}_{j_4}') \\ &= \mathcal{P}_{2,1} \mathcal{P}_{3,2} m_3(\mathbf{x}) \otimes m_2(\mathbf{x}) \end{aligned} \quad (2.69)$$

Il ne reste plus qu'à regrouper les termes en $m_3(\mathbf{x}) \otimes m_2(\mathbf{x})$, $m_3(\mathbf{x})' \otimes \text{vec} m_2(\mathbf{x})$ et $\text{vec} m_3(\mathbf{x}) \text{vec}' m_2(\mathbf{x})$ respectivement.

$$\begin{aligned} \Gamma_{(3,2)}(\mathbf{x}) &= \underbrace{(\mathbf{I}_{p^3} + \mathcal{P}_{3,2} + \mathcal{P}_{2,1} \mathcal{P}_{3,2})}_{\mathbf{R}} m_3(\mathbf{x}) \otimes m_2(\mathbf{x}) + \underbrace{(\mathbf{I}_{p^3} + \mathcal{P}_{3,2} + \mathcal{P}_{2,1} \mathcal{P}_{3,2})}_{\mathbf{R}} m_3(\mathbf{x}) \otimes m_2(\mathbf{x}) \mathbf{K}_{pp} \\ &\quad + \underbrace{(\mathbf{I}_{p^3} + \mathcal{P}_{2,1} + \mathcal{P}_{3,2} \mathcal{P}_{2,1})}_{\mathbf{P}} m_3(\mathbf{x})' \otimes \text{vec} m_2(\mathbf{x}) + \text{vec} m_3(\mathbf{x}) \text{vec}' m_2(\mathbf{x}) \\ &= \mathbf{R} m_3(\mathbf{x}) \otimes m_2(\mathbf{x}) \underbrace{(\mathbf{I}_{p^2} + \mathbf{K}_{pp})}_{\mathbf{K}} + \mathbf{P} m_3(\mathbf{x})' \otimes \text{vec} m_2(\mathbf{x}) + \text{vec} m_3(\mathbf{x}) \text{vec}' m_2(\mathbf{x}) \\ &= \mathbf{R} m_3(\mathbf{x}) \otimes m_2(\mathbf{x}) \mathbf{K} + \mathbf{P} m_3(\mathbf{x})' \otimes \text{vec} m_2(\mathbf{x}) + \text{vec} m_3(\mathbf{x}) \text{vec}' m_2(\mathbf{x}) \end{aligned} \quad (2.70)$$

2.3.6 Démonstration de l'expression à l'ordre 6

Il y a 4 partitions de 6 qui n'utilisent pas 1, ce sont (6) , $(4, 2)$, $(3, 3)$ et $(2, 2, 2)$. Ainsi $c_6(\mathbf{x})$ s'écrit comme une somme de quatre termes correspondants respectivement à ces 4 partitions.

$$c_6(\mathbf{x}) = m_6(\mathbf{x}) - \Gamma_{(4,2)}(\mathbf{x}) - \Gamma_{(3,3)}(\mathbf{x}) + 2\Gamma_{(2,2,2)}(\mathbf{x}) \quad (2.71)$$

Calcul de $\Gamma_{(4,2)}(\mathbf{x})$:

On sait que $P_{(4,2)}(\llbracket 1, 6 \rrbracket)$ possède $\frac{6!}{4!2!} = 15$ éléments.

$$\begin{aligned} P_{(4,2)}(\llbracket 1, 6 \rrbracket) = & \{ \{ \{1; 2; 3; 4\}; \{5; 6\} \}; \{ \{1; 2; 3; 5\}; \{4; 6\} \}; \{ \{1; 2; 4; 5\}; \{3; 6\} \}; \\ & \{ \{1; 3; 4; 5\}; \{2; 6\} \}; \{ \{2; 3; 4; 5\}; \{1; 6\} \}; \{ \{1; 2; 3; 6\}; \{4; 5\} \}; \\ & \{ \{1; 2; 4; 6\}; \{3; 5\} \}; \{ \{1; 3; 4; 6\}; \{2; 5\} \}; \{ \{2; 3; 4; 6\}; \{1; 5\} \}; \\ & \{ \{1; 2; 5; 6\}; \{3; 4\} \}; \{ \{1; 3; 5; 6\}; \{2; 4\} \}; \{ \{2; 3; 5; 6\}; \{1; 4\} \}; \\ & \{ \{1; 4; 5; 6\}; \{2; 3\} \}; \{ \{2; 4; 5; 6\}; \{1; 3\} \}; \{ \{3; 4; 5; 6\}; \{1; 2\} \} \end{aligned} \quad (2.72)$$

$$\Gamma_{(4,2)}(\mathbf{x}) = \sum_{\pi \in P_{(4,2)}(\llbracket 1, 6 \rrbracket)} T_{\pi}, \quad (2.73)$$

où on calcule chaque T_{π} de façon à ce que le facteur d'ordre 4 arrive en premier (on ne verra donc pas de terme en $m_2(\mathbf{x}) \otimes m_4(\mathbf{x})$ par exemple) :

$$T_{\{ \{1; 2; 3; 4\}; \{5; 6\} \}} = m_4(\mathbf{x}) \otimes m_2(\mathbf{x}) \quad (2.74)$$

$$T_{\{ \{1; 2; 3; 5\}; \{4; 6\} \}} = \mathbf{M}_4(\mathbf{x}) \otimes \text{vec}' m_2(\mathbf{x}) \quad (2.75)$$

$$T_{\{ \{1; 2; 4; 5\}; \{3; 6\} \}} = \mathcal{P}_{3,2} m_4(\mathbf{x}) \otimes m_2(\mathbf{x}) \quad (2.76)$$

$$T_{\{ \{1; 3; 4; 5\}; \{2; 6\} \}} = \mathbf{M}_4(\mathbf{x}) \otimes \text{vec}' m_2(\mathbf{x}) \mathcal{P}_{2,1} \quad (2.77)$$

$$T_{\{ \{2; 3; 4; 5\}; \{1; 6\} \}} = \mathcal{P}_{2,1} \mathcal{P}_{3,2} m_4(\mathbf{x}) \otimes m_2(\mathbf{x}) \quad (2.78)$$

$$T_{\{ \{1; 2; 3; 6\}; \{4; 5\} \}} = m_4(\mathbf{x}) \otimes m_2(\mathbf{x}) \mathcal{P}_{3,2} \quad (2.79)$$

$$T_{\{ \{1; 2; 4; 6\}; \{3; 5\} \}} = \mathbf{M}_4(\mathbf{x})' \otimes \text{vec} m_2(\mathbf{x}) \quad (2.80)$$

$$T_{\{ \{1; 3; 4; 6\}; \{2; 5\} \}} = m_4(\mathbf{x}) \otimes m_2(\mathbf{x}) \mathcal{P}_{3,2} \mathcal{P}_{2,1} \quad (2.81)$$

$$T_{\{ \{2; 3; 4; 6\}; \{1; 5\} \}} = \mathcal{P}_{2,1} \mathbf{M}_4(\mathbf{x})' \otimes \text{vec} m_2(\mathbf{x}) \quad (2.82)$$

$$T_{\{ \{1; 2; 5; 6\}; \{3; 4\} \}} = \mathcal{P}_{3,2} m_4(\mathbf{x}) \otimes m_2(\mathbf{x}) \mathcal{P}_{3,2} \quad (2.83)$$

$$T_{\{ \{1; 3; 5; 6\}; \{2; 4\} \}} = \mathbf{M}_4(\mathbf{x}) \otimes \text{vec}' m_2(\mathbf{x}) \mathcal{P}_{2,1} \mathcal{P}_{3,2} \quad (2.84)$$

$$T_{\{ \{2; 3; 5; 6\}; \{1; 4\} \}} = \mathcal{P}_{2,1} \mathcal{P}_{3,2} m_4(\mathbf{x}) \otimes m_2(\mathbf{x}) \mathcal{P}_{3,2} \quad (2.85)$$

$$T_{\{ \{1; 4; 5; 6\}; \{2; 3\} \}} = \mathcal{P}_{3,2} m_4(\mathbf{x}) \otimes m_2(\mathbf{x}) \mathcal{P}_{3,2} \mathcal{P}_{2,1} \quad (2.86)$$

$$T_{\{ \{2; 4; 5; 6\}; \{1; 3\} \}} = \mathcal{P}_{3,2} \mathcal{P}_{2,1} \mathbf{M}_4(\mathbf{x})' \otimes \text{vec} m_2(\mathbf{x}) \quad (2.87)$$

$$T_{\{ \{3; 4; 5; 6\}; \{1; 2\} \}} = \mathcal{P}_{2,1} \mathcal{P}_{3,2} m_4(\mathbf{x}) \otimes m_2(\mathbf{x}) \mathcal{P}_{3,2} \mathcal{P}_{2,1} \quad (2.88)$$

Détaillons par exemple le calcul de $T_{\{ \{1; 3; 5; 6\}; \{2; 4\} \}}$.

$$\begin{aligned} T_{\{ \{1; 3; 5; 6\}; \{2; 4\} \}} &= \sum_{(j_1, \dots, j_6) \in \llbracket 1, p \rrbracket^6} \text{mom}^{(4)}(x_{j_1}, x_{j_3}, x_{j_5}, x_{j_6}) \text{mom}^{(2)}(x_{j_2}, x_{j_4}) \mathbf{e}_{j_1} \mathbf{e}_{j_2}' \otimes \mathbf{e}_{j_3} \mathbf{e}_{j_4}' \otimes \mathbf{e}_{j_5} \mathbf{e}_{j_6}' \\ &= \sum_{(j_1, \dots, j_6) \in \llbracket 1, p \rrbracket^6} \text{mom}^{(4)}(x_{j_1}, x_{j_3}, x_{j_5}, x_{j_6}) \text{mom}^{(2)}(x_{j_2}, x_{j_4}) (\mathbf{e}_{j_1} \otimes \mathbf{e}_{j_3} \otimes \mathbf{e}_{j_5}) (\mathbf{e}_{j_2}' \otimes \mathbf{e}_{j_4}' \otimes \mathbf{e}_{j_6}') \\ &= \sum_{(j_1, \dots, j_6) \in \llbracket 1, p \rrbracket^6} \text{mom}^{(4)}(x_{j_1}, x_{j_3}, x_{j_5}, x_{j_6}) \text{mom}^{(2)}(x_{j_2}, x_{j_4}) (\mathbf{e}_{j_1} \otimes \mathbf{e}_{j_3} \otimes \mathbf{e}_{j_5}) (\mathbf{e}_{j_2}' \otimes \mathbf{e}_{j_4}' \otimes \mathbf{e}_{j_6}') \mathcal{P}_{2,1} \mathcal{P}_{3,2} \\ &= \sum_{(j_1, \dots, j_6) \in \llbracket 1, p \rrbracket^6} \text{mom}^{(4)}(x_{j_1}, x_{j_3}, x_{j_5}, x_{j_6}) \text{mom}^{(2)}(x_{j_2}, x_{j_4}) (\mathbf{e}_{j_1} \mathbf{e}_{j_6}' \otimes \mathbf{e}_{j_3} \otimes \mathbf{e}_{j_5}) \otimes (\mathbf{e}_{j_2}' \otimes \mathbf{e}_{j_4}') \mathcal{P}_{2,1} \mathcal{P}_{3,2} \\ &= \mathbf{M}_4(\mathbf{x}) \otimes \text{vec}' m_2(\mathbf{x}) \mathcal{P}_{2,1} \mathcal{P}_{3,2} \end{aligned} \quad (2.89)$$

Regroupons maintenant les termes en $m_4(\mathbf{x}) \otimes m_2(\mathbf{x})$, $\mathbf{M}_4(\mathbf{x}) \otimes \mathbf{vec}' m_2(\mathbf{x})$ et $\mathbf{M}_4(\mathbf{x})' \otimes \mathbf{vec} m_2(\mathbf{x})$ respectivement.

$$\begin{aligned}
\Gamma_{(4,2)}(\mathbf{x}) &= \underbrace{(\mathbf{I}_{p^3} + \mathcal{P}_{3,2} + \mathcal{P}_{2,1}\mathcal{P}_{3,2})}_{\mathbf{R}} m_4(\mathbf{x}) \otimes m_2(\mathbf{x}) + \underbrace{(\mathbf{I}_{p^3} + \mathcal{P}_{3,2} + \mathcal{P}_{2,1}\mathcal{P}_{3,2})}_{\mathbf{R}} m_4(\mathbf{x}) \otimes m_2(\mathbf{x}) \mathcal{P}_{3,2} \mathcal{P}_{2,1} \\
&\quad + \underbrace{(\mathbf{I}_{p^3} + \mathcal{P}_{3,2} + \mathcal{P}_{2,1}\mathcal{P}_{3,2})}_{\mathbf{R}} m_4(\mathbf{x}) \otimes m_2(\mathbf{x}) \mathcal{P}_{3,2} \\
&\quad + \mathbf{M}_4(\mathbf{x}) \otimes \mathbf{vec}' m_2(\mathbf{x}) \underbrace{(\mathbf{I}_{p^3} + \mathcal{P}_{2,1} + \mathcal{P}_{2,1}\mathcal{P}_{3,2})}_{\mathbf{P}'} \\
&\quad + \underbrace{(\mathbf{I}_{p^3} + \mathcal{P}_{2,1} + \mathcal{P}_{3,2}\mathcal{P}_{2,1})}_{\mathbf{P}} \mathbf{M}_4(\mathbf{x})' \otimes \mathbf{vec} m_2(\mathbf{x}) \\
&= \mathbf{R} \left\{ m_4(\mathbf{x}) \otimes m_2(\mathbf{x}) \right\} \mathbf{R}' + \left\{ \mathbf{M}_4(\mathbf{x}) \otimes \mathbf{vec}' m_2(\mathbf{x}) \right\} \mathbf{P}' + \mathbf{P} \left\{ \mathbf{M}_4(\mathbf{x})' \otimes \mathbf{vec} m_2(\mathbf{x}) \right\} \quad (2.90)
\end{aligned}$$

Calcul de $\Gamma_{(3,3)}(\mathbf{x})$:

On sait que $P_{(3,3)}(\llbracket 1, 6 \rrbracket)$ possède $\frac{6!}{2!3!3!} = 10$ éléments.

$$\begin{aligned}
P_{(3,3)}(\llbracket 1, 6 \rrbracket) &= \{ \{ \{1; 2; 3\}; \{4; 5; 6\} \}; \{ \{1; 2; 4\}; \{3; 5; 6\} \}; \{ \{1; 3; 4\}; \{2; 5; 6\} \}; \\
&\quad \{ \{2; 3; 4\}; \{1; 5; 6\} \}; \{ \{1; 2; 5\}; \{3; 4; 6\} \}; \{ \{1; 3; 5\}; \{2; 4; 6\} \}; \\
&\quad \{ \{2; 3; 5\}; \{1; 4; 6\} \}; \{ \{1; 4; 5\}; \{2; 3; 6\} \}; \{ \{2; 4; 5\}; \{1; 3; 6\} \}; \\
&\quad \{ \{3; 4; 5\}; \{1; 2; 6\} \} \} \quad (2.91)
\end{aligned}$$

$$\Gamma_{(3,3)}(\mathbf{x}) = \sum_{\pi \in P_{(3,3)}(\llbracket 1, 6 \rrbracket)} T_{\pi}, \quad (2.92)$$

où on calcule chaque T_{π} de façon à ce que l'éventuel facteur $m_3(\mathbf{x})$ arrive en premier (on ne verra donc pas de terme en $m_3(\mathbf{x})' \otimes m_3(\mathbf{x})$) :

$$T_{\{ \{1; 2; 3\}; \{4; 5; 6\} \}} = m_3(\mathbf{x}) \otimes m_3(\mathbf{x})' \quad (2.93)$$

$$T_{\{ \{1; 2; 4\}; \{3; 5; 6\} \}} = \mathcal{P}_{2,1} \mathcal{P}_{3,2} m_3(\mathbf{x}) \otimes m_3(\mathbf{x})' \mathcal{P}_{2,1} \mathcal{P}_{3,2} \quad (2.94)$$

$$T_{\{ \{1; 3; 4\}; \{2; 5; 6\} \}} = m_3(\mathbf{x}) \otimes m_3(\mathbf{x})' \mathcal{P}_{2,1} \quad (2.95)$$

$$T_{\{ \{2; 3; 4\}; \{1; 5; 6\} \}} = \mathcal{P}_{3,2} m_3(\mathbf{x}) \otimes m_3(\mathbf{x})' \mathcal{P}_{2,1} \mathcal{P}_{3,2} \quad (2.96)$$

$$T_{\{ \{1; 2; 5\}; \{3; 4; 6\} \}} = \mathcal{P}_{3,2} m_3(\mathbf{x}) \otimes m_3(\mathbf{x})' \quad (2.97)$$

$$T_{\{ \{1; 3; 5\}; \{2; 4; 6\} \}} = \mathbf{vec} m_3(\mathbf{x}) \otimes \mathbf{vec}' m_3(\mathbf{x}) \quad (2.98)$$

$$T_{\{ \{2; 3; 5\}; \{1; 4; 6\} \}} = \mathcal{P}_{2,1} \mathcal{P}_{3,2} m_3(\mathbf{x}) \otimes m_3(\mathbf{x})' \quad (2.99)$$

$$T_{\{ \{1; 4; 5\}; \{2; 3; 6\} \}} = \mathcal{P}_{3,2} m_3(\mathbf{x}) \otimes m_3(\mathbf{x})' \mathcal{P}_{2,1} \quad (2.100)$$

$$T_{\{ \{2; 4; 5\}; \{1; 3; 6\} \}} = m_3(\mathbf{x}) \otimes m_3(\mathbf{x})' \mathcal{P}_{2,1} \mathcal{P}_{3,2} \quad (2.101)$$

$$T_{\{ \{3; 4; 5\}; \{1; 2; 6\} \}} = \mathcal{P}_{2,1} \mathcal{P}_{3,2} m_3(\mathbf{x}) \otimes m_3(\mathbf{x})' \mathcal{P}_{2,1} \quad (2.102)$$

Il vient donc :

$$\begin{aligned}
\Gamma_{(3,3)}(\mathbf{x}) &= \underbrace{(\mathbf{I}_{p^3} + \mathcal{P}_{3,2} + \mathcal{P}_{2,1}\mathcal{P}_{3,2})}_{\mathbf{R}} m_3(\mathbf{x}) \otimes m_3(\mathbf{x})' + \underbrace{(\mathbf{I}_{p^3} + \mathcal{P}_{3,2} + \mathcal{P}_{2,1}\mathcal{P}_{3,2})}_{\mathbf{R}} m_3(\mathbf{x}) \otimes m_3(\mathbf{x})' \mathcal{P}_{2,1} \\
&\quad + \underbrace{(\mathbf{I}_{p^3} + \mathcal{P}_{3,2} + \mathcal{P}_{2,1}\mathcal{P}_{3,2})}_{\mathbf{R}} m_3(\mathbf{x}) \otimes m_3(\mathbf{x})' \mathcal{P}_{2,1} \mathcal{P}_{3,2} + \mathbf{vec} m_3(\mathbf{x}) \otimes \mathbf{vec}' m_3 \\
&= \mathbf{R} \left\{ m_3(\mathbf{x}) \otimes m_3(\mathbf{x})' \right\} \mathbf{P}' + \mathbf{vec} m_3(\mathbf{x}) \otimes \mathbf{vec}' m_3 \quad (2.103)
\end{aligned}$$

Calcul de $\Gamma_{(2,2,2)}(\mathbf{x})$:

On sait que $P_{(2,2,2)}(\llbracket 1, 6 \rrbracket)$ possède $\frac{6!}{2!2!2!} = 15$ éléments.

$$\begin{aligned}
 P_{(2,2,2)}(\llbracket 1, 6 \rrbracket) = & \{ \{1; 2\}; \{3; 4\}; \{5; 6\} \}; \{ \{1; 3\}; \{2; 4\}; \{5; 6\} \}; \{ \{2; 3\}; \{1; 4\}; \{5; 6\} \}; \\
 & \{ \{1; 2\}; \{3; 5\}; \{4; 6\} \}; \{ \{1; 3\}; \{2; 5\}; \{4; 6\} \}; \{ \{2; 3\}; \{1; 5\}; \{4; 6\} \}; \\
 & \{ \{1; 2\}; \{4; 5\}; \{3; 6\} \}; \{ \{1; 4\}; \{2; 5\}; \{3; 6\} \}; \{ \{2; 4\}; \{1; 5\}; \{3; 6\} \}; \\
 & \{ \{1; 3\}; \{4; 5\}; \{2; 6\} \}; \{ \{1; 4\}; \{3; 5\}; \{2; 6\} \}; \{ \{3; 4\}; \{1; 5\}; \{2; 6\} \}; \\
 & \{ \{2; 3\}; \{4; 5\}; \{1; 6\} \}; \{ \{2; 4\}; \{3; 5\}; \{1; 6\} \}; \{ \{3; 4\}; \{2; 5\}; \{1; 6\} \} \} \quad (2.104)
 \end{aligned}$$

$$\Gamma_{(2,2,2)}(\mathbf{x}) = \sum_{\pi \in P_{(2,2,2)}(\llbracket 1, 6 \rrbracket)} T_{\pi}, \quad (2.105)$$

où on calcule chaque T_{π} en ordonnant les facteurs dans cet ordre $m_2(\mathbf{x}) < \text{vec}m_2(\mathbf{x}) < \text{vec}'m_2(\mathbf{x})$, ce qui revient à placer en premier les blocs d'éléments (pair, impair), puis ceux (impair, impair) et enfin (pair, pair) :

$$T_{\{\{1; 2\}; \{3; 4\}; \{5; 6\}\}} = m_2(\mathbf{x}) \otimes m_2(\mathbf{x}) \otimes m_2(\mathbf{x}) \quad (2.106)$$

$$T_{\{\{1; 3\}; \{2; 4\}; \{5; 6\}\}} = \mathcal{P}_{3,2} \mathcal{P}_{2,1} m_2(\mathbf{x}) \otimes \text{vec}m_2(\mathbf{x}) \otimes \text{vec}'m_2(\mathbf{x}) \mathcal{P}_{2,1} \mathcal{P}_{3,2} \quad (2.107)$$

$$T_{\{\{2; 3\}; \{1; 4\}; \{5; 6\}\}} = \mathcal{P}_{2,1} m_2(\mathbf{x}) \otimes m_2(\mathbf{x}) \otimes m_2(\mathbf{x}) \quad (2.108)$$

$$T_{\{\{1; 2\}; \{3; 5\}; \{4; 6\}\}} = m_2(\mathbf{x}) \otimes \text{vec}m_2(\mathbf{x}) \otimes \text{vec}'m_2(\mathbf{x}) \quad (2.109)$$

$$T_{\{\{1; 3\}; \{2; 5\}; \{4; 6\}\}} = \mathcal{P}_{3,2} \mathcal{P}_{2,1} m_2(\mathbf{x}) \otimes \text{vec}m_2(\mathbf{x}) \otimes \text{vec}'m_2(\mathbf{x}) \quad (2.110)$$

$$T_{\{\{2; 3\}; \{1; 5\}; \{4; 6\}\}} = \mathcal{P}_{2,1} m_2(\mathbf{x}) \otimes \text{vec}m_2(\mathbf{x}) \otimes \text{vec}'m_2(\mathbf{x}) \quad (2.111)$$

$$T_{\{\{1; 2\}; \{4; 5\}; \{3; 6\}\}} = \mathcal{P}_{3,2} m_2(\mathbf{x}) \otimes m_2(\mathbf{x}) \otimes m_2(\mathbf{x}) \quad (2.112)$$

$$T_{\{\{1; 4\}; \{2; 5\}; \{3; 6\}\}} = \mathcal{P}_{3,2} \mathcal{P}_{2,1} m_2(\mathbf{x}) \otimes m_2(\mathbf{x}) \otimes m_2(\mathbf{x}) \quad (2.113)$$

$$T_{\{\{2; 4\}; \{1; 5\}; \{3; 6\}\}} = \mathcal{P}_{2,1} m_2(\mathbf{x}) \otimes \text{vec}m_2(\mathbf{x}) \otimes \text{vec}'m_2(\mathbf{x}) \mathcal{P}_{2,1} \mathcal{P}_{3,2} \quad (2.114)$$

$$T_{\{\{1; 3\}; \{4; 5\}; \{2; 6\}\}} = \mathcal{P}_{3,2} \mathcal{P}_{2,1} m_2(\mathbf{x}) \otimes \text{vec}m_2(\mathbf{x}) \otimes \text{vec}'m_2(\mathbf{x}) \mathcal{P}_{2,1} \quad (2.115)$$

$$T_{\{\{1; 4\}; \{3; 5\}; \{2; 6\}\}} = m_2(\mathbf{x}) \otimes \text{vec}m_2(\mathbf{x}) \otimes \text{vec}'m_2(\mathbf{x}) \mathcal{P}_{2,1} \quad (2.116)$$

$$T_{\{\{3; 4\}; \{1; 5\}; \{2; 6\}\}} = \mathcal{P}_{2,1} m_2(\mathbf{x}) \otimes \text{vec}m_2(\mathbf{x}) \otimes \text{vec}'m_2(\mathbf{x}) \mathcal{P}_{2,1} \quad (2.117)$$

$$T_{\{\{2; 3\}; \{4; 5\}; \{1; 6\}\}} = \mathcal{P}_{2,1} \mathcal{P}_{3,2} m_2(\mathbf{x}) \otimes m_2(\mathbf{x}) \otimes m_2(\mathbf{x}) \quad (2.118)$$

$$T_{\{\{2; 4\}; \{3; 5\}; \{1; 6\}\}} = m_2(\mathbf{x}) \otimes \text{vec}m_2(\mathbf{x}) \otimes \text{vec}'m_2(\mathbf{x}) \mathcal{P}_{2,1} \mathcal{P}_{3,2} \quad (2.119)$$

$$T_{\{\{3; 4\}; \{2; 5\}; \{1; 6\}\}} = \mathcal{P}_{2,1} \mathcal{P}_{3,2} \mathcal{P}_{2,1} m_2(\mathbf{x}) \otimes m_2(\mathbf{x}) \otimes m_2(\mathbf{x}) \quad (2.120)$$

Il vient donc :

$$\begin{aligned}
 \Gamma_{(2,2,2)}(\mathbf{x}) &= \underbrace{(\mathbf{I}_{p^3} + \mathcal{P}_{2,1} + \mathcal{P}_{3,2} + \mathcal{P}_{3,2} \mathcal{P}_{2,1} + \mathcal{P}_{2,1} \mathcal{P}_{3,2} + \mathcal{P}_{2,1} \mathcal{P}_{3,2} \mathcal{P}_{2,1})}_{\mathbf{Q}} m_2(\mathbf{x}) \otimes m_2(\mathbf{x}) \otimes m_2(\mathbf{x}) \\
 &\quad + \underbrace{(\mathbf{I}_{p^3} + \mathcal{P}_{2,1} + \mathcal{P}_{3,2} \mathcal{P}_{2,1})}_{\mathbf{P}} m_2(\mathbf{x}) \otimes \text{vec}m_2(\mathbf{x}) \otimes \text{vec}'m_2(\mathbf{x}) \\
 &\quad + \underbrace{(\mathbf{I}_{p^3} + \mathcal{P}_{2,1} + \mathcal{P}_{3,2} \mathcal{P}_{2,1})}_{\mathbf{P}} m_2(\mathbf{x}) \otimes \text{vec}m_2(\mathbf{x}) \otimes \text{vec}'m_2(\mathbf{x}) \mathcal{P}_{2,1} \\
 &\quad + \underbrace{(\mathbf{I}_{p^3} + \mathcal{P}_{2,1} + \mathcal{P}_{3,2} \mathcal{P}_{2,1})}_{\mathbf{P}} m_2(\mathbf{x}) \otimes \text{vec}m_2(\mathbf{x}) \otimes \text{vec}'m_2(\mathbf{x}) \mathcal{P}_{2,1} \mathcal{P}_{3,2} \\
 &= \mathbf{Q} \{ m_2(\mathbf{x}) \otimes m_2(\mathbf{x}) \otimes m_2(\mathbf{x}) \} + \mathbf{P} \{ m_2(\mathbf{x}) \otimes \text{vec}m_2(\mathbf{x}) \otimes \text{vec}'m_2(\mathbf{x}) \} \mathbf{P}' \quad (2.121)
 \end{aligned}$$

2.4 Deux conséquences faciles

2.4.1 Majoration du rang des matrices de cumulants

Corollaire 16 (rang des matrices de cumulants). *Soit \mathbf{x} un vecteur aléatoire réel de dimension $p \times 1$ vérifiant les mêmes hypothèses que dans le théorème 14. Alors les rangs de ses matrices de cumulants jusqu'à l'ordre 6 admettent les bornes ci-dessous :*

$$\text{rang}(c_2(\mathbf{x})) \leq p \quad (2.122)$$

$$\text{rang}(c_3(\mathbf{x})) \leq p \quad (2.123)$$

$$\text{rang}(c_4(\mathbf{x})) \leq \frac{p(p+1)}{2} \quad (2.124)$$

$$\text{rang}(c_5(\mathbf{x})) \leq \frac{p(p+1)}{2} \quad (2.125)$$

$$\text{rang}(c_6(\mathbf{x})) \leq \frac{p(p+1)(p+2)}{6} \quad (2.126)$$

Démonstration. Les deux premières inégalités sont évidentes. Pour les autres, posons d'abord

$$\tilde{\mathbf{K}} = \frac{1}{2}\mathbf{K} \quad \text{et} \quad \tilde{\mathbf{Q}} = \frac{1}{6}\mathbf{Q}, \quad (2.127)$$

où \mathbf{K} et \mathbf{Q} sont respectivement les matrices de tailles $p^2 \times p^2$ et $p^3 \times p^3$ données par (2.34) et (2.37) dans l'énoncé du théorème 14. Utilisant les propriétés de symétrie des cumulants, on vérifie que

$$c_4(\mathbf{x}) = \tilde{\mathbf{K}}c_4(\mathbf{x}), \quad c_5(\mathbf{x}) = c_5(\mathbf{x})\tilde{\mathbf{K}}, \quad \text{et} \quad c_6(\mathbf{x}) = \tilde{\mathbf{Q}}c_6(\mathbf{x}), \quad (2.128)$$

d'où immédiatement les majorations

$$\text{rang}(c_4(\mathbf{x})) \leq \text{rang}(\tilde{\mathbf{K}}), \quad \text{rang}(c_5(\mathbf{x})) \leq \text{rang}(\tilde{\mathbf{K}}), \quad \text{et} \quad \text{rang}(c_6(\mathbf{x})) \leq \text{rang}(\tilde{\mathbf{Q}}). \quad (2.129)$$

Or les matrices $\tilde{\mathbf{K}}$ et $\tilde{\mathbf{Q}}$ sont idempotentes (voir [Schott 2003, Theorem 1]) ; par conséquent, leur rang est donné par leur trace. Il vient donc

$$\text{rang}(\tilde{\mathbf{K}}) = \text{Tr}(\tilde{\mathbf{K}}) = \frac{1}{2}\text{Tr}(\mathbf{I}_{p^2} + \mathbf{K}_{pp}) = \frac{1}{2}(p^2 + \text{Tr}(\mathbf{K}_{pp})). \quad (2.130)$$

Par ailleurs, [Magnus 1979] ont donné la formule générale pour la trace des matrices de commutation

$$\text{Tr}(\mathbf{K}_{ab}) = 1 + \text{pgcd}(a-1, b-1) \quad (2.131)$$

En particulier, $\text{Tr}(\mathbf{K}_{pp}) = p$ et finalement

$$\text{rang}(\tilde{\mathbf{K}}) = \frac{1}{2}(p^2 + p) = \frac{p(p+1)}{2}. \quad (2.132)$$

De même, il vient

$$\begin{aligned} \text{rang}(\tilde{\mathbf{Q}}) &= \text{Tr}(\tilde{\mathbf{Q}}) = \frac{1}{6}\text{Tr}(\mathbf{Q}) \\ &= \frac{1}{6}\text{Tr}(\mathbf{I}_{p^3} + \mathbf{I}_p \otimes \mathbf{K}_{pp} + \mathbf{K}_{pp} \otimes \mathbf{I}_p + \mathbf{K}_{pp^2} + \mathbf{K}_{p^2p} + (\mathbf{K}_{pp} \otimes \mathbf{I}_p)\mathbf{K}_{p^2p}) \\ &= \frac{1}{6}(p^3 + p^2 + p^2 + p + p + p^2) \\ &= \frac{p^3 + 3p^2 + 2p}{6} = \frac{p(p+1)(p+2)}{6} \end{aligned} \quad (2.133)$$

□

2.4.2 Matrices des moments d'ordres 4 et 6 d'un vecteur gaussien

Les expressions des matrices des moments d'ordre 4 et 6 d'un vecteur gaussien ont été un objectif de plusieurs chercheurs : elles ont été données respectivement par [Magnus 1979], [Neudecker 1969], [Tracy 1993] et récemment [Schott 2003]. Ce résultat peut être obtenu ici comme simple corollaire du théorème 14.

Corollaire 17 (Matrice des moments d'un vecteur gaussien). *Soit $\mathbf{z} \sim \mathcal{N}(\mathbf{0}_p, \Sigma)$, un vecteur gaussien centré. Alors ses matrices des moments d'ordres 4 et 6 sont*

$$m_4(\mathbf{z}) = \mathbf{K} \left\{ \Sigma \otimes \Sigma \right\} + \text{vec} \Sigma \text{vec}' \Sigma \quad (2.134)$$

$$m_6(\mathbf{z}) = \mathbb{E}(\mathbf{z}\mathbf{z}' \otimes \mathbf{z}\mathbf{z}' \otimes \mathbf{z}\mathbf{z}') = \mathbf{Q} (\Sigma \otimes \Sigma \otimes \Sigma) + \mathbf{P} (\Sigma \otimes \text{vec} \Sigma \otimes \text{vec}' \Sigma) \mathbf{P}' \quad (2.135)$$

où les matrices \mathbf{K}, \mathbf{Q} et \mathbf{P} sont celles du théorème 14.

Démonstration. Comme \mathbf{z} est gaussien, ses cumulants d'ordres > 2 sont nuls, d'où (2.134). Nous utilisons la formule (2.19) pour $k = 6$, dans laquelle les seuls termes non nuls seront ceux correspondant à une partition de type $(2, 2, 2)$. On a donc :

$$m_6(\mathbf{z}) = \Gamma_{(2,2,2)}(\mathbf{x}), \quad (2.136)$$

où ici

$$\Gamma_{(2,2,2)}(\mathbf{x}) = \sum_{\pi \in P_{(2,2,2)}(\llbracket 1,6 \rrbracket)} \left(\sum_{(j_1, \dots, j_6) \in \llbracket 1, p \rrbracket^6} \prod_{B \in \pi} \text{cum}^{(2)}(x_{j_l} : l \in B) \mathbf{e}_{j_1} \mathbf{e}_{j_2}' \otimes \mathbf{e}_{j_3} \mathbf{e}_{j_4}' \otimes \mathbf{e}_{j_5} \mathbf{e}_{j_6}' \right), \quad (2.137)$$

quantité qui a déjà été calculée en (2.121), puisque $m_2(\mathbf{x}) = c_2(\mathbf{x}) = \Sigma$. \square

2.5 Gain en temps de calcul pour l'estimation des cumulants

2.5.1 Estimation des matrices de moments et cumulants

D'après ce qui précède, la matrice de cumulant d'ordre $k \leq 6$, peut être estimée empiriquement à partir des matrices des moments d'ordres inférieurs en s'appuyant sur les formules matricielles du théorème (14). En effet, soit \mathbf{x} de dimension p , échantillonné sur une longueur d'échantillonnage T , sous forme d'une matrice

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}(1) & \mathbf{x}(2) & \dots & \mathbf{x}(T) \end{pmatrix} \quad (2.138)$$

alors, les matrices des moments estimées sont :

$$\hat{m}_2(\mathbf{x}) = \frac{1}{T} \sum_{i=1}^T \mathbf{x}(i) \mathbf{x}(i)' \quad (2.139)$$

$$\hat{m}_3(\mathbf{x}) = \frac{1}{T} \sum_{i=1}^T \mathbf{x}(i) \mathbf{x}(i)' \otimes \mathbf{x}(i) \quad (2.140)$$

$$\hat{m}_4(\mathbf{x}) = \frac{1}{T} \sum_{i=1}^T \mathbf{x}(i) \mathbf{x}(i)' \otimes \mathbf{x}(i) \mathbf{x}(i)' \quad (2.141)$$

$$\hat{m}_5(\mathbf{x}) = \frac{1}{T} \sum_{i=1}^T \mathbf{x}(i) \mathbf{x}(i)' \otimes \mathbf{x}(i) \mathbf{x}(i)' \otimes \mathbf{x}(i) \quad (2.142)$$

$$\hat{m}_6(\mathbf{x}) = \frac{1}{T} \sum_{i=1}^T \mathbf{x}(i) \mathbf{x}(i)' \otimes \mathbf{x}(i) \mathbf{x}(i)' \otimes \mathbf{x}(i) \mathbf{x}(i)' \quad (2.143)$$

De même, la matrice $\mathbf{M}_4(\mathbf{x})$, impliquée dans la formule de $c_6(\mathbf{x})$ est estimée par

$$\widehat{\mathbf{M}}_4(\mathbf{x}) = \frac{1}{T} \sum_{i=1}^T \mathbf{x}(i) \mathbf{x}(i)' \otimes \mathbf{x}(i) \otimes \mathbf{x}(i) \quad (2.144)$$

2.5.2 Comparaisons numériques entre les estimateurs matriciels et les estimateurs scalaires de cumulants

Pour monter l'avantage de cette étude d'un point de vue pratique, considérons sur Matlab un vecteur $\mathbf{x} = \text{rand}(p, 1000)$. Nous allons faire une simple comparaison entre les temps qu'il faut pour estimer la matrices $c_5(\mathbf{x})$ (respectivement, $c_6(\mathbf{x})$), en utilisant dans un premier temps leurs formules scalaires [Mac Cullagh 1987] (voir les expressions scalaires à l'annexe A), puis un rangement réalisé via la commande Matlab `reshape`(\mathcal{T}_5, p^3, p^2) (respectivement `reshape`(\mathcal{T}_6, p^3, p^3)) (où \mathcal{T}_k est le tenseur de cumulant d'ordre k voir [Albera 2004]); dans un second temps nous estimons directement les matrices de cumulant à partir de leurs expressions matricielles du théorème 14. Les tests ont été réalisés sur une machine Samsung (Processeur : Intel(R) Core(TM) i3-2310M CPU 2.10GHZ, Mémoire(RAM) 4 Go) et la comparaison est faite par rapport à la dimension p du vecteur. Notons que dans cet exercice, nous n'avons pas utilisé l'asymétrie de ces quantités dans la procédure d'estimation ni pour les expressions scalaire ni pour les matricielles.

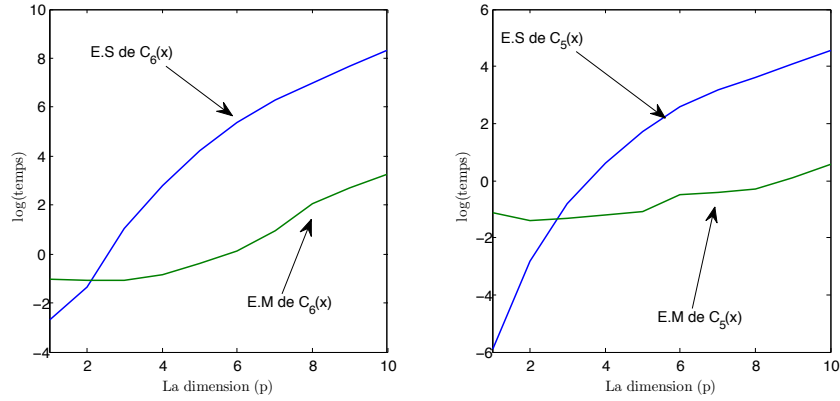


FIGURE 2.1 – Comparaisons entre les temps d'exécution (en seconde) d'une estimation de la matrice de cumulants d'ordre cinq à droite (resp. six à gauche). En bleu, une estimation basée sur l'expression scalaire (E.S) scalaire et en vert, une estimation basée sur l'expression matricielle (E.M), pour une longueur de signal fixé à 1000 échantillons, en fonction de la dimension p du vecteur \mathbf{x} , qui varie de 1 à 10.

2.6 Conclusion

Dans ce chapitre, nous avons proposé des nouvelles expressions matricielles (ordre 5 et 6) des matrices de cumulants. Ces expressions ont plusieurs avantages que ce soit au point de vue théorique ou numérique, car elles permettent immédiatement de donner des résultats sur les rangs de ces matrices (voir le lemme 16). Nous verrons plus loin dans cette thèse que le rang d'une matrice de cumulants est un indicateur très important qui caractérise la dépendance. En plus, comme nous l'avons constaté dans la comparaison numérique (voir la section (2.5.2), la figure (2.1)), ces expressions permettent une estimation ultra-rapide de ces quantités dans les langages de programmation matriciels comme Matlab.

Analyse en composantes indépendantes multidimensionnelles via des cumulants d'ordres variés

Résumé : Nous montrons que sous les hypothèses usuelles de l'analyse en composantes indépendantes multidimensionnelles (ACIM), les matrices de cumulants contractées à un ordre statistique r (quelconque) sont toutes bloc-diagonalisables dans la même base. Nous en déduisons des algorithmes résolution d'ACIM par bloc-diagonalisation conjointe et comparons les résultats obtenus aux ordres 3 à 6, entre eux et avec d'autres méthodes, sur quelques signaux synthétiques.

3.1 Introduction et motivations

La plupart des algorithmes initialement proposés pour l'analyse en composantes indépendantes multidimensionnelles (ACIM) étaient conçus pour résoudre uniquement le cas particulier de l'ACIM de taille fixe. C'est seulement après la conception des algorithmes de bloc-diagonalisation conjointe (BDC) pouvant traiter le cas où les blocs diagonaux n'ont pas (nécessairement) la même taille [Theis 2007a, Maehara 2011, Murota 2010], que des chercheurs sont parvenus à traiter le cas général en bloc-diagonalisant des matrices issues des statistiques d'ordre deux [Theis 2005a, H. Gutch 2010, Lahat 2009, Lahat 2012a], issues de la fonction caractéristique ou de la densité de probabilité [Theis 2005b, Theis 2005a], ou encore des matrices de dispersion [Nordhausen 2011].

Toutefois, les solutions d'ACIM exploitant les cumulants d'ordre supérieur à deux, qui offrent plus d'information que les cumulants d'ordre deux, surtout lorsque certaines sources ne possèdent pas de diversité temporelle, restent rares. Les auteurs de [Castella 2007] ont étudié les cumulants d'ordres trois et quatre de sources dépendantes en considérant plusieurs cas de figures, puis ont montré que leurs cumulants peuvent avoir des propriétés intéressantes. Dans [Theis 2007a], un algorithme de bloc diagonalisation conjointe baptisé *JBD by JD* a été intégré dans la méthode d'ACI ordinaire JADE [Cardoso 1993], pour engendrer une solution du problème d'ACIM. Sans formuler théoriquement le problème d'ACIM sous forme d'un problème de BDC, l'auteur propose de remplacer simplement l'étape de diagonalisation conjointe de JADE par une étape de BDC. L'algorithme résultant a été baptisé SJADE (pour Subspace-JADE). Toutefois, les questions suivantes restent ouvertes :

- Les matrices de cumulants contractées d'ordre quatre (notion définie initialement dans [Cardoso 1993] et reprise dans [Theis 2007a]) des observations sont-elles bloc-diagonalisables dans la base \mathbf{A} ?
- Plus généralement, les matrices de cumulants contractées d'ordre r (notion qui sera définie ici dans la section 3.4) sont-elles bloc-diagonalisables dans la base \mathbf{A} ?

Autrement dit : le problème d'ACIM peut-il être reformulé à un ordre statistique r quelconque, comme un problème de bloc-diagonalisation conjointe (BDC) ?

À l'ordre deux, la réponse est positive et bien connue [Theis 2005a, H. Gutch 2010]. Aux ordres supérieurs,

la réponse est loin d'être aussi simple et fait l'un des objets de ce chapitre. Nous commençons par rappeler les hypothèses usuelles qui sous-tendent le modèle ACIM (§3.2). Nous rappelons ensuite les définitions des matrices de cumulant (§3.3) puis des matrices de cumulant contractées (§3.4). Le cœur de ce chapitre fait l'objet de la section §3.5 où nous montrons que les matrices de cumulants contractées de tout ordre sont toutes bloc-diagonalisables dans la même base. Ce résultat est exploité à la section §3.6 où nous présentons une famille d'algorithmes baptisés SJADE_{r₁,...,r_q} qui généralise le désormais classique algorithme JADE au cas multidimensionnel et à des ordres statistiques arbitraires. Ces algorithmes sont comparés à la section §3.9.

3.2 Simplification du modèle ACIM

Nous cherchons p composantes indépendantes multidimensionnelles, de dimension $n_i = \dim(\mathbf{s}_i)$, $i \in \llbracket 1, p \rrbracket$. En général, les dimensions ne sont pas connues à l'avance car elles sont inobservables directement à partir des observations. Dans cette étude, nous nous plaçons dans un cadre général où les dimensions des sources ne sont pas (nécessairement) égales ni connues d'avance.

3.2.1 Sur l'unicité aux indéterminations près

Etant donné un modèle ACIM des observations \mathbf{x} comme en (1.1), si on se donne des matrices inversibles \mathbf{D}_i de taille $n_i \times n_i$ pour $i \in \llbracket 1, p \rrbracket$, et une permutation $\sigma \in \mathfrak{S}_p$, alors on a encore

$$\begin{pmatrix} \mathbf{A}_{\sigma(1)}\mathbf{D}_{\sigma(1)} & \cdots & \mathbf{A}_{\sigma(p)}\mathbf{D}_{\sigma(p)} \end{pmatrix} \begin{pmatrix} \mathbf{D}_{\sigma(1)}^{-1}\mathbf{s}_{\sigma(1)} \\ \vdots \\ \mathbf{D}_{\sigma(p)}^{-1}\mathbf{s}_{\sigma(p)} \end{pmatrix} = \mathbf{x}_{\sigma(1)} + \cdots + \mathbf{x}_{\sigma(p)} = \mathbf{x} \quad (3.1)$$

Ainsi, les composantes indépendantes \mathbf{x}_i ne sont déterminées qu'à permutation près (si celle-ci n'échange entre elles que des composantes de même dimension); tandis que les sources ne sont déterminées qu'à permutation et matrice inversible près. On dira que \mathbf{B} est une matrice de séparation du problème d'ACIM si $\mathbf{BA} = \mathbf{PD}$, où \mathbf{P} est une matrice permutation et $\mathbf{D} = \text{bdiag}(\mathbf{D}_1, \mathbf{D}_2, \dots, \mathbf{D}_p)$ est une matrice bloc-diagonale.

3.2.2 Hypothèses

Pour garantir l'unicité du modèle (aux indéterminations triviales près comme nous venons de le voir), nous devons adopter les hypothèses suivantes :

- (h₁) la matrice \mathbf{A} est de rang n .
- (h₂) au-plus une source indépendante (scalaire) peut vérifier que le cumulant d'ordre r est nul.
- (h₃) les sources indépendantes sont irréductibles.

L'irréductibilité des sources est fondamentale comme l'ont expliqué [Theis 2007a, Gutch 2007, Gutch 2012]. Rappelons simplement qu'un vecteur aléatoire \mathbf{x} de dimension n est irréductible s'il est impossible de l'écrire

$$\mathbf{x} = \mathbf{A} \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix}, \quad (3.2)$$

où \mathbf{A} matrice inversible, \mathbf{y}_1 et \mathbf{y}_2 vecteurs aléatoires indépendants, de dimensions respectives $k \neq 0$ et $n - k$ (voir [Gutch 2012]).

3.2.3 Le blanchiment d'ordre deux

Sans perte de généralité, nous supposons que le vecteur source est centré ($\mathbb{E}(\mathbf{s}) = \mathbf{0}$) et réduit ($\mathbf{R}_s = \mathbb{E}(\mathbf{s}\mathbf{s}') = \mathbf{I}_n$), (voir par exemple [Cardoso 1998, Theis 2007a]); cela implique que $\mathbf{R}_x = \mathbb{E}(\mathbf{x}\mathbf{x}') = \mathbf{A}\mathbf{A}'$. La matrice de covariance \mathbf{R}_x des observations $\mathbf{x}_{m \times 1}$, qui est symétrique et semi-définie positive, admet une décomposition en valeurs singulières $\mathbf{R}_x = \mathbf{U}\mathbf{D}\mathbf{U}'$, où \mathbf{D} est une matrice diagonale $n \times n$, n étant le nombre de valeurs singulières non nulles (et donc strictement positives); $\mathbf{U}_{m \times n}$ est une matrice de taille $m \times n$ vérifiant $\mathbf{U}'\mathbf{U} = \mathbf{I}_n$.

Définissons la matrice de blanchiment $\mathbf{W}_{n \times m} = \mathbf{D}^{-\frac{1}{2}}\mathbf{U}'$ et posons $\tilde{\mathbf{x}} = (\mathbf{W}\mathbf{x})_{n \times 1}$; il vient alors

$$\mathbf{R}_{\tilde{\mathbf{x}}} = \mathbb{E}(\tilde{\mathbf{x}}\tilde{\mathbf{x}}') = \mathbf{W}\mathbf{R}_x\mathbf{W}' = \mathbf{D}^{-\frac{1}{2}}\mathbf{U}'(\mathbf{U}\mathbf{D}\mathbf{U}')\mathbf{U}\mathbf{D}^{-\frac{1}{2}} = \mathbf{I}_n \quad (3.3)$$

Puisque $\mathbf{R}_x = \mathbf{A}\mathbf{A}'$, on voit aussi que $\mathbf{W}\mathbf{A}\mathbf{A}'\mathbf{W}' = \mathbf{I}_n$; autrement dit la matrice $\tilde{\mathbf{A}} = \mathbf{W}\mathbf{A}$, qui est de taille $n \times n$, est orthogonale. Par multiplication par la matrice de blanchiment, le modèle (1.1) se transforme en $\tilde{\mathbf{x}} = \tilde{\mathbf{A}}\mathbf{s}$; c'est pourquoi nous supposerons directement le modèle (1.1) tel que la matrice \mathbf{A} soit orthogonale (donc en particulier, $m = n$). Résoudre le problème d'ACIM consiste alors à déterminer, via les statistiques d'ordre supérieurs à deux, une matrice orthogonale \mathbf{A} .

3.2.4 Notations

La matrice de mélange \mathbf{A} peut être partagée en matrices colonnes \mathbf{A}_i , de taille $n \times n_i$ (voir le modèle (1.1)), ou en vecteurs lignes que nous notons ℓ_j :

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \dots & \mathbf{A}_p \end{pmatrix} = \begin{pmatrix} \ell_1 \\ \ell_2 \\ \vdots \\ \ell_n \end{pmatrix} \quad (3.4)$$

La matrice \mathbf{A}_i peut elle-même être partagée en lignes par

$$\mathbf{A}_i = \begin{pmatrix} \ell_1^i \\ \ell_2^i \\ \vdots \\ \ell_n^i \end{pmatrix} \quad (3.5)$$

où $\forall k \in \llbracket 1, n \rrbracket$, ℓ_k^i , de taille $1 \times n_i$, est le k -ème vecteur ligne de \mathbf{A}_i . De même, chaque ligne de la matrice de mélange \mathbf{A} peut être exprimée en fonction de lignes des matrices \mathbf{A}_i , de la façon suivante

$$\ell_k = \begin{pmatrix} \ell_k^1 & \ell_k^2 & \dots & \ell_k^p \end{pmatrix} \quad (3.6)$$

3.3 Matrices de cumulants

Notre outil de base pour résoudre le problème d'ACIM sera les cumulants. Il est bien connu que les cumulants scalaires peuvent être calculés à partir de l'évaluation des moments scalaires, via la formule classique de Leonov et Shiryaev (voir par exemple [Mac Cullagh 1987] et [Albera 2002] qui attribue une nature tensorielle aux cumulants). Néanmoins, il est souvent souhaitable de ranger les cumulants dans des matrices, carrées symétriques lorsqu'il s'agit d'un ordre pair, ce qui se réalise en pratique en utilisant la fonction `reshape` de Matlab (voir [Albera 2004]). Ce rangement facilite en effet l'exploitation de propriétés algébriques telles que la redondance matricielle (voir [Albera 2003, Albera 2005, Albera 2004]), ou la diagonalisation (voir [Kollo 2008]).

Définition 8 (matrice de cumulants¹). Soit \mathbf{x} un vecteur aléatoire de dimension $n \times 1$. Alors, la matrice de cumulants d'ordre r , est

$$c_r(\mathbf{x}) = \sum_{i_1, i_2, \dots, i_r=1}^n \text{cum}(x_{i_1}, x_{i_2}, \dots, x_{i_r}) \mathbf{E}_{i_1, i_2, \dots, i_r} \quad (3.7)$$

où

$$\mathbf{E}_{i_1, i_2, \dots, i_r} = \begin{cases} \left(\mathbf{e}_{i_1} \otimes \mathbf{e}_{i_3} \otimes \dots \otimes \mathbf{e}_{i_{2k-1}} \right) \left(\mathbf{e}_{i_2} \otimes \mathbf{e}_{i_4} \otimes \dots \otimes \mathbf{e}_{i_{2k}} \right)' & \text{si } r = 2k \\ \left(\mathbf{e}_{i_1} \otimes \mathbf{e}_{i_3} \otimes \dots \otimes \mathbf{e}_{i_{2k+1}} \right) \left(\mathbf{e}_{i_2} \otimes \mathbf{e}_{i_4} \otimes \dots \otimes \mathbf{e}_{i_{2k}} \right)' & \text{si } r = 2k + 1 \end{cases} \quad (3.8)$$

où $(\mathbf{e}_i)_{i \in \llbracket 1, n \rrbracket}$ désigne la base canonique de \mathbb{R}^n et \otimes désigne le produit de Kronecker.

Signalons que des expressions matricielles de $c_r(\mathbf{x})$ en fonction des cumulants d'ordres inférieurs à r permettent sur Matlab une estimation plus rapide de ces quantités (voir le chapitre 2). Les formules pour les ordres 3 et 4 sont données dans [Kollo 2008], tandis que le chapitre 2 contient les expressions jusqu'à l'ordre 6. La définition 8 accorde à une matrice de cumulants d'ordre pair ($r = 2k$) une forme carrée symétrique de taille $n^k \times n^k$; par exemple à l'ordre $r = 4$, nous aurons $c_4(\mathbf{x})$, matrice de taille $n^2 \times n^2$ appelée aussi quadricovariance [Albera 2005]; à l'ordre $r = 6$, nous aurons $c_6(\mathbf{x})$, de taille $n^3 \times n^3$ appelée hexacovariance [Albera 2003]. Pour l'ordre impair $r = 2k + 1$, la matrice de cumulants sera de taille $n^{k+1} \times n^k$. Plusieurs propriétés intéressantes des matrices de cumulants peuvent être déduites directement de la définition 8 ci-dessus. Nous en regroupons quelques-unes dans la proposition suivante.

Proposition 18. (i) si \mathbf{z} est un vecteur aléatoire gaussien, alors

$$\forall r > 2, \quad c_r(\mathbf{z}) = \mathbf{0}. \quad (3.9)$$

(ii) si \mathbf{x} et \mathbf{y} sont deux vecteurs aléatoires indépendants de même dimension, alors

$$\forall r \in \mathbb{N} \quad c_r(\mathbf{x} + \mathbf{y}) = c_r(\mathbf{x}) + c_r(\mathbf{y}) \quad (3.10)$$

(iii) si $\mathbf{x} = \mathbf{A}\mathbf{s}$, où \mathbf{A} est une matrice de taille $m \times n$ et \mathbf{x}, \mathbf{s} sont des vecteurs aléatoires de tailles respectives $m \times 1$ et $n \times 1$, alors

$$c_r(\mathbf{x}) = \begin{cases} \underbrace{(\mathbf{A} \otimes \dots \otimes \mathbf{A})}_{k \text{ fois}} c_r(\mathbf{s}) \underbrace{(\mathbf{A} \otimes \dots \otimes \mathbf{A})'}_{k \text{ fois}} = (\otimes^k \mathbf{A}) c_r(\mathbf{s}) (\otimes^k \mathbf{A}') & \text{si } r = 2k \\ \underbrace{(\mathbf{A} \otimes \dots \otimes \mathbf{A})}_{k+1 \text{ fois}} c_r(\mathbf{s}) \underbrace{(\mathbf{A} \otimes \dots \otimes \mathbf{A})'}_{k \text{ fois}} = (\otimes^{k+1} \mathbf{A}) c_r(\mathbf{s}) (\otimes^k \mathbf{A}') & \text{si } r = 2k + 1 \end{cases} \quad (3.11)$$

(iv) si $\mathbf{x} = \mathbf{A}\mathbf{s} = \begin{pmatrix} \mathbf{A}_1 & \dots & \mathbf{A}_p \end{pmatrix} \begin{pmatrix} \mathbf{s}_1 \\ \vdots \\ \mathbf{s}_p \end{pmatrix} \in \mathbb{R}^m$, où les $\mathbf{s}_i \in \mathbb{R}^{n_i}$ sont des vecteurs aléatoires

1. Cette définition est rien d'autre que la définition 8 donnée au niveau du chapitre 2, nous avons uniquement changer les notations pour une question de simplicité.

indépendants, $\mathbf{A}_i \in \mathbb{R}^{m \times n_i}$, $n = n_1 + \dots + n_p$ et $\mathbf{A} \in \mathbb{R}^{m \times n}$, alors pour tout entier r ,

$$c_r(\mathbf{x}) = \begin{cases} \mathcal{A}_k \begin{pmatrix} c_r(\mathbf{s}_1) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & c_r(\mathbf{s}_p) \end{pmatrix} \mathcal{A}'_k & \text{si } r = 2k \\ \mathcal{A}_{k+1} \begin{pmatrix} c_r(\mathbf{s}_1) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & c_r(\mathbf{s}_p) \end{pmatrix} \mathcal{A}'_k & \text{si } r = 2k + 1 \end{cases} \quad (3.12)$$

où les matrices \mathcal{A}_k sont données par

$$\mathcal{A}_k = (\otimes^k \mathbf{A}_1 \quad \otimes^k \mathbf{A}_2 \quad \cdots \quad \otimes^k \mathbf{A}_p) \in \mathbb{R}^{m^k \times \sum_{j=1}^p n_j^k} \quad (3.13)$$

Démonstration. (i) découle du fait que $\forall r > 2$, $\text{cum}(z_1, z_2, \dots, z_r) = 0$, par définition d'un vecteur gaussien.

(ii) découle du fait bien connu (sinon voir par exemple [Mendel 1991, p. 280]) que

$$\text{cum}(x_1 + y_1, x_2 + y_2, \dots, x_r + y_r) = \text{cum}(x_1, x_2, \dots, x_r) + \text{cum}(y_1, y_2, \dots, y_r) \quad (3.14)$$

lorsque les vecteurs aléatoires \mathbf{x} et \mathbf{y} sont indépendants.

(iii) Nous montrons seulement le cas $r = 2k$, le cas $r = 2k + 1$ pouvant se traiter de manière analogue. Il suffit d'utiliser la multilinéarité des cumulants, en écrivant $x_j = \sum_{i=1}^n a_{j,i} s_i$ pour $j \in \llbracket 1, m \rrbracket$, où $\mathbf{A} = (a_{j,i})_{(j,i) \in \llbracket 1, m \rrbracket \times \llbracket 1, n \rrbracket}$.

$$\begin{aligned} c_r(\mathbf{x}) &= \sum_{j_1, j_2, \dots, j_{2k}=1}^m \text{cum}(x_{j_1}, \dots, x_{j_{2k}}) \mathbf{E}_{j_1, j_2, \dots, j_{2k}} \\ &= \sum_{j_1, j_2, \dots, j_{2k}=1}^m \text{cum} \left(\sum_{i_1=1}^n a_{j_1, i_1} s_{i_1}, \dots, \sum_{i_{2k}=1}^n a_{j_{2k}, i_{2k}} s_{i_{2k}} \right) \mathbf{E}_{j_1, j_2, \dots, j_{2k}} \\ &= \sum_{j_1, j_2, \dots, j_{2k}=1}^m \left(\sum_{i_1, i_2, \dots, i_{2k}=1}^n a_{j_1, i_1} \cdots a_{j_{2k}, i_{2k}} \text{cum}(s_{i_1}, \dots, s_{i_{2k}}) \right) \mathbf{E}_{j_1, j_2, \dots, j_{2k}} \\ &= \sum_{i_1, i_2, \dots, i_{2k}=1}^n \text{cum}(s_{i_1}, \dots, s_{i_{2k}}) \left(\sum_{j_1, j_2, \dots, j_{2k}=1}^m a_{j_1, i_1} \cdots a_{j_{2k}, i_{2k}} \mathbf{E}_{j_1, j_2, \dots, j_{2k}} \right) \\ &= \sum_{i_1, i_2, \dots, i_{2k}=1}^n \text{cum}(s_{i_1}, \dots, s_{i_{2k}}) \sum_{j_1, j_2, \dots, j_{2k}=1}^m a_{j_1, i_1} \cdots a_{j_{2k}, i_{2k}} (\mathbf{e}_{j_1} \mathbf{e}'_{j_2} \otimes \cdots \otimes \mathbf{e}_{j_{2k-1}} \mathbf{e}'_{j_{2k}}) \end{aligned} \quad (3.15)$$

Ci-dessus, $(\mathbf{e}_j)_{j \in \llbracket 1, m \rrbracket}$ désigne la base canonique de \mathbb{R}^m . Si nous désignons par $(\mathbf{f}_i)_{i \in \llbracket 1, n \rrbracket}$ la base canonique

de \mathbb{R}^n , il vient $\mathbf{A}\mathbf{f}_i = \sum_{j=1}^m a_{j,i} \mathbf{e}_j$, pour tout $i \in \llbracket 1, n \rrbracket$; par suite,

$$\begin{aligned}
 c_r(\mathbf{x}) &= \sum_{i_1, i_2, \dots, i_{2k}=1}^n \text{cum}(s_{i_1}, \dots, s_{i_{2k}}) \left(\left(\sum_{j_1=1}^m a_{j_1, i_1} \mathbf{e}_{j_1} \right) \left(\sum_{j_2=1}^m a_{j_2, i_2} \mathbf{e}_{j_2} \right)' \otimes \dots \right. \\
 &\quad \left. \dots \otimes \left(\sum_{j_{2k-1}=1}^m a_{j_{2k-1}, i_{2k-1}} \mathbf{e}_{j_{2k-1}} \right) \left(\sum_{j_{2k}=1}^m a_{j_{2k}, i_{2k}} \mathbf{e}_{j_{2k}} \right)' \right) \\
 &= \sum_{i_1, i_2, \dots, i_{2k}=1}^n \text{cum}(s_{i_1}, \dots, s_{i_{2k}}) ((\mathbf{A}\mathbf{f}_{i_1}) (\mathbf{A}\mathbf{f}_{i_2})' \otimes \dots \otimes (\mathbf{A}\mathbf{f}_{i_{2k-1}}) (\mathbf{A}\mathbf{f}_{i_{2k}})') \\
 &= \sum_{i_1, i_2, \dots, i_{2k}=1}^n \text{cum}(s_{i_1}, \dots, s_{i_{2k}}) (\mathbf{A} \otimes \dots \otimes \mathbf{A}) (\mathbf{f}_{i_1} \mathbf{f}_{i_2}' \otimes \dots \otimes \mathbf{f}_{i_{2k-1}} \mathbf{f}_{i_{2k}}') (\mathbf{A} \otimes \dots \otimes \mathbf{A})' \\
 &= (\mathbf{A} \otimes \dots \otimes \mathbf{A}) \left(\sum_{i_1, i_2, \dots, i_{2k}=1}^n \text{cum}(s_{i_1}, \dots, s_{i_{2k}}) (\mathbf{f}_{i_1} \mathbf{f}_{i_2}' \otimes \dots \otimes \mathbf{f}_{i_{2k-1}} \mathbf{f}_{i_{2k}}') \right) (\mathbf{A} \otimes \dots \otimes \mathbf{A})' \\
 &= (\mathbf{A} \otimes \dots \otimes \mathbf{A}) c_r(\mathbf{s}) (\mathbf{A} \otimes \dots \otimes \mathbf{A})' \tag{3.16}
 \end{aligned}$$

(iv) Encore une fois, nous ne montrons que le cas $r = 2k$, l'autre se traitant de manière analogue. Nous avons

$$\mathbf{x} = \mathbf{A}\mathbf{s} = \sum_{j=1}^p \mathbf{A}_j \mathbf{s}_j \tag{3.17}$$

L'indépendance des \mathbf{s}_j implique celle des $\mathbf{A}_j \mathbf{s}_j$. Utilisant successivement (ii) et (iii), il vient

$$\begin{aligned}
 c_r(\mathbf{x}) &= c_r(\mathbf{A}\mathbf{s}) = c_r \left(\sum_{j=1}^p \mathbf{A}_j \mathbf{s}_j \right) \\
 &= \sum_{j=1}^p c_r(\mathbf{A}_j \mathbf{s}_j) \\
 &= \sum_{j=1}^p \otimes^k \mathbf{A}_j c_r(\mathbf{s}_j) \otimes^k \mathbf{A}_j' \\
 &= \underbrace{(\otimes^k \mathbf{A}_1 \quad \dots \quad \otimes^k \mathbf{A}_p)}_{=\mathcal{A}_k} \begin{pmatrix} c_r(\mathbf{s}_1) & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & c_r(\mathbf{s}_p) \end{pmatrix} \underbrace{\begin{pmatrix} \otimes^k \mathbf{A}_1' \\ \vdots \\ \otimes^k \mathbf{A}_p' \end{pmatrix}}_{=\mathcal{A}_k'} \tag{3.18}
 \end{aligned}$$

□

3.4 Matrice de cumulant contractée

Cardoso et Soulomiac [Cardoso 1993] développent une approche matricielle pour l'ACI basée sur les cumulants d'ordre quatre connue sous le nom de JADE. Ils définissent, à l'aide du tenseur de cumulant d'ordre quatre, la notion de matrice de cumulant "contractée" (nous avons ajouté ici l'adjectif contractée pour ne pas confondre avec la matrice de cumulants $c_4(\mathbf{x})$ de la définition 8) :

Définition 9 (matrice de cumulants contractée d'ordre 4). Soient $\mathbf{M} = (m_{i_1, i_2})_{(i_1, i_2) \in \llbracket 1, n \rrbracket^2}$ une matrice de taille $n \times n$ et \mathbf{x} un vecteur aléatoire de taille $n \times 1$. La matrice de cumulant contractée d'ordre quatre est la matrice de taille $n \times n$, notée $\mathcal{Q}_{(4)}^{\mathbf{x}}[\mathbf{M}]$, dont le terme général est

$$\left(\mathcal{Q}_{(4)}^{\mathbf{x}}[\mathbf{M}]\right)_{i_3, i_4} = \sum_{i_1, i_2=1}^n \text{cum}(x_{i_1}, x_{i_2}, x_{i_3}, x_{i_4}) m_{i_1, i_2}. \quad (3.19)$$

Le produit matriciel étoile, défini par [MacRae 1974] (voir les rappels en B.3), permet d'exprimer la matrice de cumulant contractée de façon globale, directement à partir de la matrice contractante et de la matrice de cumulant.

Lemme 19. Soit $\mathbf{M} = (m_{i_1, i_2})_{(i_1, i_2) \in \llbracket 1, n \rrbracket^2}$ une matrice de taille $n \times n$; alors

$$\mathcal{Q}_{(4)}^{\mathbf{x}}[\mathbf{M}] = \mathbf{M} * c_4(\mathbf{x}) = \sum_{i_1, i_2=1}^n m_{i_1, i_2} [c_4(\mathbf{x})]_{i_1, i_2} \quad (3.20)$$

où $[c_4(\mathbf{x})]_{i_1, i_2}$ est la matrice de taille $n \times n$ correspondant au bloc (i_1, i_2) de la matrice $c_4(\mathbf{x})$.

Démonstration. Il suffit de remarquer que l'élément (i_3, i_4) de la matrice $[c_4(\mathbf{x})]_{i_1, i_2}$ est justement

$$\left([c_4(\mathbf{x})]_{i_1, i_2}\right)_{i_3, i_4} = \text{cum}(x_{i_1}, x_{i_2}, x_{i_3}, x_{i_4}) \quad (3.21)$$

On peut démontrer ce résultat aussi en appliquant les propriétés (P1) à (P3) (voir l'annexe B.18, B.19 et B.20) du produit étoile $*$ à l'expression de $c_4(\mathbf{x})$ (voir théorème 14), nous obtiendrons facilement que

$$\mathbf{M} * c_4(\mathbf{x}) = \mathbb{E}(\mathbf{x}' \mathbf{M} \mathbf{x} \mathbf{x} \mathbf{x}') - \mathbf{R}_{\mathbf{x}} \mathbf{M} \mathbf{R}_{\mathbf{x}} - \mathbf{R}_{\mathbf{x}} \mathbf{M}' \mathbf{R}_{\mathbf{x}} - \text{Tr}(\mathbf{R}_{\mathbf{x}} \mathbf{M}) \mathbf{R}_{\mathbf{x}} \quad (3.22)$$

qui est l'expression de $\mathcal{Q}_{(4)}^{\mathbf{x}}[\mathbf{M}]$ (voir [Cardoso 1999]). \square

Afin de généraliser la notion de matrice de cumulant contractée à un ordre r quelconque, nous allons utiliser les matrices $\mathbf{E}_{i_1, i_2, \dots, i_q}$ déjà définies en (3.8).

Définition 10 (matrice de cumulants contractée d'ordre r quelconque). Pour $q \in \mathbb{N}$, désignons par $\mathbf{M}^{(q)}$ une matrice écrite sous la forme

$$\mathbf{M}^{(q)} = \sum_{i_1, i_2, \dots, i_q=1}^n m_{i_1, i_2, \dots, i_q} \mathbf{E}_{i_1, i_2, \dots, i_q} \quad (3.23)$$

où les matrices $\mathbf{E}_{i_1, i_2, \dots, i_q}$ sont celles déjà définies en (3.8) et m_{i_1, i_2, \dots, i_q} des constantes arbitraires. Alors, on définit la matrice de cumulants contractée d'ordre r et de dimension $n \times n$, par

$$\mathcal{Q}_{(r)}^{\mathbf{x}}[\mathbf{M}^{(r-2)}] = \mathbf{M}^{(r-2)} * c_r(\mathbf{x}) = \sum_{i_1, i_2, \dots, i_{r-2}=1}^n m_{i_1, i_2, \dots, i_{r-2}} [c_r(\mathbf{x})]_{i_1, i_2, \dots, i_{r-2}} \quad (3.24)$$

où $[c_r(\mathbf{x})]_{i_1, i_2, \dots, i_{r-2}}$ désigne la matrice repérée par $(i_1, i_2, \dots, i_{r-2})$ dans le partage de $c_r(\mathbf{x})$ en matrices de taille $n \times n$; autrement dit, l'élément (i_{r-1}, i_r) de la matrice $[c_r(\mathbf{x})]_{i_1, i_2, \dots, i_{r-2}}$ est

$$\left([c_r(\mathbf{x})]_{i_1, i_2, \dots, i_{r-2}}\right)_{i_{r-1}, i_r} = \text{cum}(x_{i_1} x_{i_2} \cdots x_{i_r}). \quad (3.25)$$

Remarque 20. Evidemment, il s'ensuit que l'élément (i_{r-1}, i_r) de la matrice de cumulants contractée est

$$\left(\mathcal{Q}_{(r)}^{\mathbf{x}}[\mathbf{M}^{(r-2)}]\right)_{i_{r-1}, i_r} = \sum_{i_1, i_2, \dots, i_{r-2}=1}^n \text{cum}(x_{i_1} x_{i_2} \cdots x_{i_r}) m_{i_1, i_2, \dots, i_{r-2}}, \quad (3.26)$$

ce qui correspond à la généralisation naturelle de (3.19).

3.5 Formulation du problème ACIM à l'ordre r comme problème de BDC

A partir du modèle $\mathbf{x} = \mathbf{A}\mathbf{s}$ comme en (1.1), l'indépendance statistique des sources \mathbf{s}_i implique, d'après la propriété (iv) de la proposition 18, que

$$c_r(\mathbf{x}) = \begin{cases} \mathcal{A}_k \mathcal{C}_{2k}(\mathbf{s}) \mathcal{A}'_k & \text{si } r = 2k \\ \mathcal{A}_{k+1} \mathcal{C}_{2k+1}(\mathbf{s}) \mathcal{A}'_k & \text{si } r = 2k + 1 \end{cases} \quad (3.27)$$

où nous avons posé

$$\mathcal{C}_r(\mathbf{s}) = \text{bdiag} \left(c_r(\mathbf{s}_1), c_r(\mathbf{s}_2), \dots, c_r(\mathbf{s}_p) \right) \quad (3.28)$$

Avant d'annoncer le résultat principal de ce chapitre, qui fera l'objet du théorème 22, précisons la structure des matrices \mathcal{A}_k .

Lemme 21 (structure des matrices \mathcal{A}_k).

Pour tout entier $k \geq 1$, la matrice $\mathcal{A}_k = \left(\otimes^k \mathbf{A}_1 \quad \otimes^k \mathbf{A}_2 \quad \dots \quad \otimes^k \mathbf{A}_p \right)$, de taille $n^k \times \sum_{j=1}^p n_j^k$, s'écrit comme une seule colonne matrice de la manière suivante

$$\mathcal{A}_k = \begin{pmatrix} \mathbf{A}\mathbf{D}_{11\dots 11} \\ \mathbf{A}\mathbf{D}_{11\dots 12} \\ \vdots \\ \mathbf{A}\mathbf{D}_{11\dots 1n} \\ \mathbf{A}\mathbf{D}_{11\dots 21} \\ \mathbf{A}\mathbf{D}_{11\dots 22} \\ \vdots \\ \mathbf{A}\mathbf{D}_{11\dots 2n} \\ \vdots \\ \mathbf{A}\mathbf{D}_{1n\dots nn} \\ \mathbf{A}\mathbf{D}_{2n\dots nn} \\ \vdots \\ \mathbf{A}\mathbf{D}_{nn\dots nn} \end{pmatrix} \quad (3.29)$$

où, pour tout $(i_1, i_2, \dots, i_{k-1}) \in \llbracket 1, n \rrbracket^{k-1}$, la matrice $\mathbf{D}_{i_1 i_2 \dots i_{k-1}}$ est de taille $n \times \sum_{j=1}^p n_j^k$ et s'exprime par

$$\mathbf{D}_{i_1 i_2 \dots i_{k-1}} = \text{bdiag} \left(\ell_{i_1}^1 \otimes \dots \otimes \ell_{i_{k-1}}^1 \otimes \mathbf{I}_{n_1}, \ell_{i_1}^2 \otimes \dots \otimes \ell_{i_{k-1}}^2 \otimes \mathbf{I}_{n_2}, \dots, \ell_{i_1}^p \otimes \dots \otimes \ell_{i_{k-1}}^p \otimes \mathbf{I}_{n_p} \right) \quad (3.30)$$

où on a désigné par ℓ_k^i la k -ème ligne de la matrice \mathbf{A}_i (voir les notations en sous-section 3.2.4).

La matrice \mathcal{A}_k est appelé la k -ème puissance de Khatri-Rao [Kagan 1973].

Démonstration. Pour simplifier les notations, nous allons montrer le lemme 21, uniquement pour $k \in \{2, 3\}$, une démonstration par récurrence pouvant être déduite de ces deux cas. Par ailleurs, ce chapitre n'utilise que ces deux cas.

Le cas $k = 2$. Nous avons

$$\mathcal{A}_2 = \begin{pmatrix} \mathbf{A}_1 \otimes \mathbf{A}_1 & \mathbf{A}_2 \otimes \mathbf{A}_2 & \dots & \mathbf{A}_p \otimes \mathbf{A}_p \end{pmatrix} = \begin{pmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \\ \vdots \\ \mathbf{B}_n \end{pmatrix} \quad (3.31)$$

où, pour $i \in \llbracket 1, n \rrbracket$, la matrice \mathbf{B}_i est de taille $n \times \sum_{j=1}^p n_j^2$. Elle s'écrit

$$\mathbf{B}_i = \begin{pmatrix} \ell_i^1 \otimes \mathbf{A}_1 & \ell_i^2 \otimes \mathbf{A}_2 & \dots & \ell_i^p \otimes \mathbf{A}_p \end{pmatrix}, \quad (3.32)$$

où ℓ_i^j , de taille $1 \times n_j$, est la i -ème ligne de la matrice \mathbf{A}_j (voir (3.6)). D'après les propriétés élémentaires du produit de Kronecker, il vient

$$\ell_i^j \otimes \mathbf{A}_j = (1 \ell_i^j) \otimes (\mathbf{A}_j \mathbf{I}_{n_j}) = \mathbf{A}_j (\ell_i^j \otimes \mathbf{I}_{n_j}) \quad (3.33)$$

En utilisant (3.33) dans (3.32), nous obtenons

$$\begin{aligned} \mathbf{B}_i &= \begin{pmatrix} \mathbf{A}_1(\ell_i^1 \otimes \mathbf{I}_{n_1}) & \mathbf{A}_2(\ell_i^2 \otimes \mathbf{I}_{n_2}) & \dots & \mathbf{A}_p(\ell_i^p \otimes \mathbf{I}_{n_p}) \end{pmatrix} \\ &= \underbrace{\begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \dots & \mathbf{A}_p \end{pmatrix}}_{=\mathbf{A}} \underbrace{\begin{pmatrix} \ell_i^1 \otimes \mathbf{I}_{n_1} & \mathbf{0}_{n_1 \times n_2^2} & \dots & \mathbf{0}_{n_1 \times n_p^2} \\ \mathbf{0}_{n_2 \times n_1^2} & \ell_i^2 \otimes \mathbf{I}_{n_2} & \dots & \mathbf{0}_{n_2 \times n_p^2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{n_p \times n_1^2} & \mathbf{0}_{n_p \times n_2^2} & \dots & \ell_i^p \otimes \mathbf{I}_{n_p} \end{pmatrix}}_{=\mathbf{D}_i} \\ &= \mathbf{A} \text{bdiag}(\ell_i^1 \otimes \mathbf{I}_{n_1}, \ell_i^2 \otimes \mathbf{I}_{n_2}, \dots, \ell_i^p \otimes \mathbf{I}_{n_p}) \end{aligned} \quad (3.34)$$

Le cas $k = 3$.

$$\begin{aligned} \mathcal{A}_3 &= \begin{pmatrix} \mathbf{A}_1 \otimes \mathbf{A}_1 \otimes \mathbf{A}_1 & \mathbf{A}_2 \otimes \mathbf{A}_2 \otimes \mathbf{A}_2 & \dots & \mathbf{A}_p \otimes \mathbf{A}_p \otimes \mathbf{A}_p \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{B}_{11} \\ \mathbf{B}_{12} \\ \vdots \\ \mathbf{B}_{1n} \\ \hline \mathbf{B}_{21} \\ \mathbf{B}_{22} \\ \vdots \\ \mathbf{B}_{2n} \\ \hline \vdots \\ \hline \mathbf{B}_{n1} \\ \mathbf{B}_{n2} \\ \vdots \\ \mathbf{B}_{nn} \end{pmatrix} = \begin{pmatrix} \mathcal{B}_1 \\ \mathcal{B}_2 \\ \vdots \\ \mathcal{B}_n \end{pmatrix} \end{aligned} \quad (3.35)$$

Les matrices \mathbf{B}_{ij} ont été choisies de même taille $n \times \sum_{k=1}^p n_k^3$. Chaque matrice \mathcal{B}_i est de taille $n^2 \times \sum_{k=1}^p n_k^3$

et s'écrit

$$\begin{aligned}
 \mathcal{B}_i &= \begin{pmatrix} \mathbf{B}_{i1} \\ \mathbf{B}_{i2} \\ \vdots \\ \mathbf{B}_{in} \end{pmatrix} = \begin{pmatrix} \ell_i^1 \otimes \mathbf{A}_1 \otimes \mathbf{A}_1 & \ell_i^2 \otimes \mathbf{A}_2 \otimes \mathbf{A}_2 & \cdots & \ell_i^p \otimes \mathbf{A}_p \otimes \mathbf{A}_p \end{pmatrix} \\
 &= \begin{pmatrix} (\mathbf{A}_1 \otimes \mathbf{A}_1)(\ell_i^1 \otimes \mathbf{I}_{n_1^2}) & \cdots & (\mathbf{A}_p \otimes \mathbf{A}_p)(\ell_i^p \otimes \mathbf{I}_{n_p^2}) \end{pmatrix} \\
 &= \mathcal{A}_2 \text{ bdiag} \left(\ell_i^1 \otimes \mathbf{I}_{n_1^2}, \ell_i^2 \otimes \mathbf{I}_{n_2^2}, \cdots, \ell_i^p \otimes \mathbf{I}_{n_p^2} \right)
 \end{aligned} \tag{3.36}$$

Remplaçant \mathcal{A}_2 par sa valeur (voir le cas $k = 2$), nous obtenons

$$\mathcal{B}_i = \begin{pmatrix} \mathbf{A}\mathbf{D}_1 \\ \mathbf{A}\mathbf{D}_2 \\ \vdots \\ \mathbf{A}\mathbf{D}_n \end{pmatrix} \text{ bdiag} \left(\ell_i^1 \otimes \mathbf{I}_{n_1^2}, \ell_i^2 \otimes \mathbf{I}_{n_2^2}, \cdots, \ell_i^p \otimes \mathbf{I}_{n_p^2} \right), \tag{3.37}$$

ce qui veut dire que

$$\begin{aligned}
 \mathbf{B}_{ij} &= \mathbf{A}\mathbf{D}_j \text{ bdiag} \left(\ell_i^1 \otimes \mathbf{I}_{n_1^2}, \ell_i^2 \otimes \mathbf{I}_{n_2^2}, \cdots, \ell_i^p \otimes \mathbf{I}_{n_p^2} \right) \\
 &= \mathbf{A} \text{ bdiag} \left(\ell_j^1 \otimes \mathbf{I}_{n_1}, \cdots, \ell_j^p \otimes \mathbf{I}_{n_p} \right) \text{ bdiag} \left(\ell_i^1 \otimes \mathbf{I}_{n_1^2}, \cdots, \ell_i^p \otimes \mathbf{I}_{n_p^2} \right) \\
 &= \mathbf{A} \text{ bdiag} \left(1 \otimes \ell_j^1 \otimes \mathbf{I}_{n_1}, \cdots, 1 \otimes \ell_j^p \otimes \mathbf{I}_{n_p} \right) \text{ bdiag} \left(\ell_i^1 \otimes \mathbf{I}_{n_1} \otimes \mathbf{I}_{n_1}, \cdots, \ell_i^p \otimes \mathbf{I}_{n_p} \otimes \mathbf{I}_{n_p} \right) \\
 &= \underbrace{\mathbf{A} \text{ bdiag} \left(\ell_i^1 \otimes \ell_j^1 \otimes \mathbf{I}_{n_1}, \ell_i^2 \otimes \ell_j^2 \otimes \mathbf{I}_{n_2}, \cdots, \ell_i^p \otimes \ell_j^p \otimes \mathbf{I}_{n_p} \right)}_{=\mathbf{D}_{ij}}
 \end{aligned} \tag{3.38}$$

□

De ce lemme 21 et de l'équation (3.27) on déduit directement que la matrice de cumulants $c_r(\mathbf{x})$ peut aussi être vue comme une matrice partagée en blocs de taille $n \times n$ qui sont donnés par

$$\left[c_r(\mathbf{x}) \right]_{i_1, i_2, \dots, i_{r-2}} = \begin{cases} \mathbf{A}\mathbf{D}_{i_1 i_3 \dots i_{r-3}} \mathcal{C}_{2k}(\mathbf{s}) \mathbf{D}'_{i_2 i_4 \dots i_{r-2}} \mathbf{A}' & \text{si } r = 2k \\ \mathbf{A}\mathbf{D}_{i_1 i_3 \dots i_{r-2}} \mathcal{C}_{2k+1}(\mathbf{s}) \mathbf{D}'_{i_2 i_4 \dots i_{r-3}} \mathbf{A}' & \text{si } r = 2k + 1 \end{cases} \tag{3.39}$$

Nous sommes maintenant en mesure d'énoncer notre résultat.

Théorème 22 (bloc diagonalisation conjointe des matrices de cumulants contractées).

Soit $\mathbf{s} = (\mathbf{s}'_1, \mathbf{s}'_2, \dots, \mathbf{s}'_p)'$ un vecteur aléatoire où les \mathbf{s}_i sont des vecteurs aléatoires de dimension $\dim(\mathbf{s}_i) = n_i$ qu'on suppose indépendants. On pose $n = \sum_{i=1}^p n_i$. Soit \mathbf{A} une matrice de taille $n \times n$. On pose $\mathbf{x} = \mathbf{A}\mathbf{s}$. Alors, pour tout entier $r > 2$, nous avons :

$$\mathcal{Q}_{(r)}^{\mathbf{x}}[\mathbf{M}^{(r-2)}] = \mathbf{A} \text{ bdiag} \left(\Delta_{\mathbf{s}_1}^{(r)}, \Delta_{\mathbf{s}_2}^{(r)}, \cdots, \Delta_{\mathbf{s}_p}^{(r)} \right) \mathbf{A}' \tag{3.40}$$

où

$$(\forall i \in \llbracket 1, p \rrbracket) \quad \Delta_{\mathbf{s}_i}^{(r)} = \begin{cases} \mathcal{Q}_{(2k)}^{\mathbf{s}_i} \left[\left(\otimes^{k-1} \mathbf{A}'_i \right) \mathbf{M}^{(2k-2)} \left(\otimes^{k-1} \mathbf{A}_i \right) \right] & \text{si } r = 2k \\ \mathcal{Q}_{(2k+1)}^{\mathbf{s}_i} \left[\left(\otimes^k \mathbf{A}'_i \right) \mathbf{M}^{(2k-1)} \left(\otimes^{k-1} \mathbf{A}_i \right) \right] & \text{si } r = 2k + 1 \end{cases} \tag{3.41}$$

Démonstration. Nous nous restreindrons encore à l'ordre pair $r = 2k$, la démonstration du théorème pour l'ordre impair $r = 2k + 1$ utilisant la même démarche.

Dans $c_{2k}(\mathbf{x}) = \mathcal{A}_k \mathcal{C}_{2k}(\mathbf{s}) \mathcal{A}'_k$, remplaçons \mathcal{A}_k par sa valeur donnée au lemme 21. Nous voyons alors que le bloc de taille $n \times n$ de la matrice $c_{2k}(\mathbf{x})$ indexé par $(i_1, j_1, \dots, i_{k-1}, j_{k-1})$ s'écrit

$$\left[c_{2k}(\mathbf{x}) \right]_{i_1, j_1, \dots, i_{k-1}, j_{k-1}} = \mathbf{A} \mathbf{D}_{i_1 \dots i_{k-1}} \mathcal{C}_{2k}(\mathbf{s}) \mathbf{D}'_{j_1 \dots j_{k-1}} \mathbf{A}' \quad (3.42)$$

Avec la définition (3.24) de la matrice de cumulants contractée d'ordre $2k$, il vient

$$\begin{aligned} \mathcal{Q}_{(2k)}^{\mathbf{x}}[\mathbf{M}] &= \sum_{i_1, j_1, \dots, i_{k-1}, j_{k-1}=1}^n m_{i_1, j_1, \dots, i_{k-1}, j_{k-1}} \left[c_{2k}(\mathbf{x}) \right]_{i_1, j_1, \dots, i_{k-1}, j_{k-1}} \\ &= \sum_{i_1, j_1, \dots, i_{k-1}, j_{k-1}=1}^n m_{i_1, j_1, \dots, i_{k-1}, j_{k-1}} \mathbf{A} \mathbf{D}_{i_1 \dots i_{k-1}} \mathcal{C}_{2k}(\mathbf{s}) \mathbf{D}'_{j_1 \dots j_{k-1}} \mathbf{A}' \\ &= \mathbf{A} \underbrace{\left(\sum_{i_1, j_1, \dots, i_{k-1}, j_{k-1}=1}^n m_{i_1, j_1, \dots, i_{k-1}, j_{k-1}} \mathbf{D}_{i_1 \dots i_{k-1}} \mathcal{C}_{2k}(\mathbf{s}) \mathbf{D}'_{j_1 \dots j_{k-1}} \right)}_{=\Delta} \mathbf{A}' \end{aligned} \quad (3.43)$$

On voit que Δ est bloc-diagonale, puisque $\mathbf{D}_{i_1 \dots i_{k-1}}$, $\mathcal{C}_{2k}(\mathbf{s})$ et $\mathbf{D}'_{j_1 \dots j_{k-1}}$ le sont.

$$\Delta = \sum_{i_1, j_1, \dots, i_{k-1}, j_{k-1}=1}^n m_{i_1, j_1, \dots, i_{k-1}, j_{k-1}} \mathbf{D}_{i_1 \dots i_{k-1}} \mathcal{C}_{2k}(\mathbf{s}) \mathbf{D}'_{j_1 \dots j_{k-1}} = \text{bdiag} \left(\Delta_{\mathbf{s}_1}^{(2k)}, \dots, \Delta_{\mathbf{s}_p}^{(2k)} \right), \quad (3.44)$$

où, en utilisant (3.28) et (3.30), nous avons

$$\Delta_{\mathbf{s}_i}^{(2k)} = \sum_{i_1, j_1, \dots, i_{k-1}, j_{k-1}=1}^n m_{i_1, j_1, \dots, i_{k-1}, j_{k-1}} \left(\ell_{i_1}^i \otimes \dots \otimes \ell_{i_{k-1}}^i \otimes \mathbf{I}_{n_i} \right) c_{2k}(\mathbf{s}_i) \left(\ell_{j_1}^i \otimes \dots \otimes \ell_{j_{k-1}}^i \otimes \mathbf{I}_{n_i} \right)' \quad (3.45)$$

En utilisant maintenant la propriété (B.21) du produit étoile, nous obtenons

$$\begin{aligned} \Delta_{\mathbf{s}_i}^{(2k)} &= \left(\sum_{i_1, j_1, \dots, i_{k-1}, j_{k-1}=1}^n m_{i_1, j_1, \dots, i_{k-1}, j_{k-1}} \left(\ell_{i_1}^i \otimes \dots \otimes \ell_{i_{k-1}}^i \right)' \left(\ell_{j_1}^i \otimes \dots \otimes \ell_{j_{k-1}}^i \right) \right) * c_{2k}(\mathbf{s}_i) \\ &= \left(\sum_{i_1, j_1, \dots, i_{k-1}, j_{k-1}=1}^n m_{i_1, j_1, \dots, i_{k-1}, j_{k-1}} \left(\mathbf{A}'_i \mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{A}'_i \mathbf{e}_{i_{k-1}} \right) \left(\mathbf{e}'_{j_1} \mathbf{A}_i \otimes \dots \otimes \mathbf{e}'_{j_1} \mathbf{A}_i \right) \right) * c_{2k}(\mathbf{s}_i) \\ &= \left(\sum_{i_1, j_1, \dots, i_{k-1}, j_{k-1}=1}^n m_{i_1, j_1, \dots, i_{k-1}, j_{k-1}} \left(\otimes^{k-1} \mathbf{A}'_i \right) \left(\mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{e}_{i_{k-1}} \right) \left(\mathbf{e}'_{j_1} \otimes \dots \otimes \mathbf{e}'_{j_1} \right) \left(\otimes^{k-1} \mathbf{A}_i \right) \right) * c_{2k}(\mathbf{s}_i) \\ &= \left(\left(\otimes^{k-1} \mathbf{A}'_i \right) \left(\sum_{i_1, j_1, \dots, i_{k-1}, j_{k-1}=1}^n m_{i_1, j_1, \dots, i_{k-1}, j_{k-1}} \mathbf{E}_{i_1, j_1, \dots, i_{k-1}, j_{k-1}} \right) \left(\otimes^{k-1} \mathbf{A}_i \right) \right) * c_{2k}(\mathbf{s}_i) \\ &= \left(\left(\otimes^{k-1} \mathbf{A}'_i \right) \mathbf{M}^{(2k-2)} \left(\otimes^{k-1} \mathbf{A}_i \right) \right) * c_{2k}(\mathbf{s}_i) \\ &= \mathcal{Q}_{(2k)}^{\mathbf{s}_i} \left[\left(\otimes^{k-1} \mathbf{A}'_i \right) \mathbf{M}^{(2k-2)} \left(\otimes^{k-1} \mathbf{A}_i \right) \right] \end{aligned} \quad (3.46)$$

□

Corollaire 23 (Le cas ACI). *Supposons toutes les sources scalaires; autrement dit, $p = n$ et $n_i = 1$ pour tout $i \in \llbracket 1, p \rrbracket$. Alors, en notant par \mathbf{a}_i les colonnes de \mathbf{A} et par $k_i^{(r)} = c_r(\mathbf{s}_i) = \text{cum}(\mathbf{s}_i, \dots, \mathbf{s}_i)$ le cumulante scalaire d'ordre r de la source \mathbf{s}_i , nous avons*

$$\mathcal{Q}_{(r)}^{\mathbf{x}}[\mathbf{M}^{(r-2)}] = \mathbf{A} \text{diag} \left(\delta_{\mathbf{s}_1}^{(r)}, \delta_{\mathbf{s}_2}^{(r)}, \dots, \delta_{\mathbf{s}_p}^{(r)} \right) \mathbf{A}' \quad (3.47)$$

où

$$(\forall i \in \llbracket 1, n \rrbracket) \quad \delta_{\mathbf{s}_i}^{(r)} = \begin{cases} ((\otimes^{k-1} \mathbf{a}'_i) \mathbf{M}^{(2k-2)} (\otimes^{k-1} \mathbf{a}_i)) k_i^{(2k)} & \text{si } r = 2k \\ ((\otimes^k \mathbf{a}'_i) \mathbf{M}^{(2k-1)} (\otimes^{k-1} \mathbf{a}_i)) k_i^{(2k+1)} & \text{si } r = 2k + 1 \end{cases} \quad (3.48)$$

Le théorème 22 signifie que toute matrice de cumulants contractée $\mathcal{Q}_{(r)}^{\mathbf{x}}[\mathbf{M}]$ de dimension $n \times n$ est bloc diagonalisable dans la base \mathbf{A} , et que le i -ème bloc diagonal de cette matrice est la matrice de cumulants contractée $\Delta_{\mathbf{s}_i}^{(r)}$ de taille $n_i \times n_i$. En particulier, lorsque toutes les sources sont scalaires (c'est-à-dire ACIM=ACI), et si de plus $2k = 4$, le corollaire 23 retrouve le résultat classique

$$\mathcal{Q}_{(4)}^{\mathbf{x}}[\mathbf{M}^{(2)}] = \mathbf{A} \text{diag} \left(k_1^{(4)} \mathbf{a}'_1 \mathbf{M}^{(2)} \mathbf{a}_1, k_2^{(4)} \mathbf{a}'_2 \mathbf{M}^{(2)} \mathbf{a}_2, \dots, k_p^{(4)} \mathbf{a}'_p \mathbf{M}^{(2)} \mathbf{a}_p \right) \mathbf{A}' \quad (3.49)$$

où $k_i^{(4)}$ s'appelle dans ce cas le kurtosis de la source \mathbf{s}_i . Par conséquent, il est clair que le corollaire 23 étend à l'ordre r le résultat obtenu auparavant dans [Cardoso 1993, page 5] pour l'ACI d'ordre 4 et que le théorème 22 est sa généralisation multidimensionnelle. Dans le cadre ACIM, explicitons les résultats du théorème 22 pour les ordres $r \in \{3; 4; 5; 6\}$ (les seuls utilisés dans les simulations numériques présentées plus loin) :

$$\mathcal{Q}_{(3)}^{\mathbf{x}}[\mathbf{M}^{(1)}] = \mathbf{A} \text{bdiag} \left(\mathcal{Q}_{(3)}^{\mathbf{s}_1} [\mathbf{A}'_1 \mathbf{M}^{(1)}], \dots, \mathcal{Q}_{(3)}^{\mathbf{s}_p} [\mathbf{A}'_p \mathbf{M}^{(1)}] \right) \mathbf{A}' \quad (3.50)$$

$$\mathcal{Q}_{(4)}^{\mathbf{x}}[\mathbf{M}^{(2)}] = \mathbf{A} \text{bdiag} \left(\mathcal{Q}_{(4)}^{\mathbf{s}_1} [\mathbf{A}'_1 \mathbf{M}^{(2)} \mathbf{A}_1], \dots, \mathcal{Q}_{(4)}^{\mathbf{s}_p} [\mathbf{A}'_p \mathbf{M}^{(2)} \mathbf{A}_p] \right) \mathbf{A}' \quad (3.51)$$

$$\mathcal{Q}_{(5)}^{\mathbf{x}}[\mathbf{M}^{(3)}] = \mathbf{A} \text{bdiag} \left(\mathcal{Q}_{(5)}^{\mathbf{s}_1} [(\mathbf{A}'_1 \otimes \mathbf{A}'_1) \mathbf{M}^{(3)} \mathbf{A}_1], \dots, \mathcal{Q}_{(5)}^{\mathbf{s}_p} [(\mathbf{A}'_p \otimes \mathbf{A}'_p) \mathbf{M}^{(3)} \mathbf{A}_p] \right) \mathbf{A}' \quad (3.52)$$

$$\mathcal{Q}_{(6)}^{\mathbf{x}}[\mathbf{M}^{(4)}] = \mathbf{A} \text{bdiag} \left(\mathcal{Q}_{(6)}^{\mathbf{s}_1} [(\mathbf{A}'_1 \otimes \mathbf{A}'_1) \mathbf{M}^{(4)} (\mathbf{A}_1 \otimes \mathbf{A}_1)], \dots, \mathcal{Q}_{(6)}^{\mathbf{s}_p} [(\mathbf{A}'_p \otimes \mathbf{A}'_p) \mathbf{M}^{(4)} (\mathbf{A}_p \otimes \mathbf{A}_p)] \right) \mathbf{A}' \quad (3.53)$$

avec $\mathbf{M}^{(1)} : n \times 1$, $\mathbf{M}^{(2)} : n \times n$, $\mathbf{M}^{(3)} : n^2 \times n$ et $\mathbf{M}^{(4)} : n^2 \times n^2$.

3.6 SJADE_r : une famille d'algorithmes pour l'ACIM

Le résultat du théorème 22 nous permet de proposer une extension multidimensionnelle et à tout ordre de l'algorithme JADE proposé par [Cardoso 1993]. En effet le théorème 22 montre que si les observations \mathbf{x} sont générées selon le modèle (1.1) avec les sources \mathbf{s}_i indépendantes, alors les matrices de cumulants contractées $\mathcal{Q}_{(r)}^{\mathbf{x}}[\mathbf{M}^{(r-2)}]$ sont toutes bloc-diagonales dans la même base \mathbf{A} . D'où l'espoir de récupérer la matrice de mélange et les dimensions des sources en bloc-diagonalisant conjointement (BDC) un ensemble de matrices de cumulants contractées.

Il existe maintenant deux familles d'algorithme de BDC :

- (i) les algorithmes de BDC orthogonale, qui estiment une matrice orthogonale qui BDC l'ensemble de matrices en question; c'est le cas des algorithmes *JBD* ([Abed-Meraim 2004, Févotte 2007a]), *JBD by JD* ([Theis 2007a]), *MKKKM* ([Murota 2010, Maehara 2010]) ainsi que sa version améliorée *commdec* ([Maehara 2011]);
- (ii) les algorithmes dont la matrice de bloc-diagonalisation n'est plus supposée orthogonale, qui ont été conçu spécialement pour éviter l'étape de blanchiment d'ordre deux [Nion 2011, Ghennioui 2010, Ghennioui 2007].

Un choix naturel pour l'ensemble de matrices contractantes consiste à choisir $\mathbf{M}^{(r-2)} = \mathbf{E}_{i_1, i_2, \dots, i_{r-2}}$, ce qui implique $\mathcal{Q}_{(r)}^{\mathbf{x}}[\mathbf{E}_{i_1, i_2, \dots, i_{r-2}}] = [c_r(\mathbf{x})]_{i_1, i_2, \dots, i_{r-2}}$. Remarquons que l'extension multidimensionnelle à l'ordre quatre coïncide avec l'algorithme baptisé SJADE qui est proposé dans [Theis 2007a]. Voici maintenant le synoptique de l'algorithme baptisé SJADE_r, qui résout le problème d'ACIM à l'ordre statistique r .

Algorithme 24 (SJADE_r).

1. **blanchiment des observations :**

- (a) $\mathbf{x} = \mathbf{x} - \mathbb{E}(\mathbf{x})$ (centrage) ;
- (b) **estimation de la matrice de blanchiment :** calcul de la matrice de covariance $\mathbf{R}_{\mathbf{x}} = \mathbb{E}(\mathbf{x}\mathbf{x}')$; décomposition en valeurs singulières de $\mathbf{R}_{\mathbf{x}}$; estimation du nombre n de sources (c'est le nombre de valeurs propres non nulles) et obtention de la matrice de blanchiment $\mathbf{W} : n \times m$;
- (c) $\mathbf{x} = \mathbf{W}\mathbf{x}$ (blanchiment) ;

2. **construction d'un ensemble de matrices de cumulants contractées :** par exemple l'ensemble

$$\mathcal{M}^{(r)} = \left\{ \mathcal{Q}_{(r)}^{\mathbf{x}}[\mathbf{E}_{i_1, i_2, \dots, i_{r-2}}] \quad ; \quad 1 \leq i_1, i_2, \dots, i_{r-2} \leq n \right\} \quad (3.54)$$

de, a priori, n^{r-2} matrices, calculé à partir d'une simple estimation empirique de la matrice de cumulants $c_r(\mathbf{x})$. Cette étape utilise les formules développées dans ([Ould-Baba 2013] le chapitre 2).

3. **bloc-diagonalisation conjointe (BDC) :** estimation de la matrice orthogonale \mathbf{A}' qui bloc-diagonalise conjointement $\mathcal{M}^{(r)}$.

4. **estimation des sources :**

- (a) $\mathbf{B} = \mathbf{A}'\mathbf{W}$ (matrice de séparation) ;
- (b) $\mathbf{s} = \mathbf{B}\mathbf{x}$.

Remarque 25 (Réduction du nombre des matrices à bloc-diagonaliser). *Compte-tenu des symétries de la matrice de cumulants $c_r(\mathbf{x})$, l'ensemble*

$$\left\{ \mathbf{E}_{i_1, i_2, \dots, i_{r-2}} \quad ; \quad 1 \leq i_1, i_2, \dots, i_{r-2} \leq n \right\}$$

des matrices contractantes est très redondant. Précisément, sur l'expression (3.26) des matrices de cumulants contractées, nous voyons que pour toute permutation $\sigma \in \mathfrak{S}_{r-2}$,

$$\begin{aligned} \left(\mathcal{Q}_{(r)}^{\mathbf{x}}[\sigma \cdot \mathbf{M}^{(r-2)}] \right)_{i_{r-1}, i_r} &= \sum_{i_1, i_2, \dots, i_{r-2}=1}^n \underbrace{\text{cum}(x_{i_1} x_{i_2} \dots x_{i_r})}_{=\text{cum}(x_{i_{\sigma(1)}} \dots x_{i_{\sigma(r-2)}} x_{i_{r-1}} x_{i_r})} m_{i_{\sigma(1)}, i_{\sigma(2)}, \dots, i_{\sigma(r-2)}} \\ &= \left(\mathcal{Q}_{(r)}^{\mathbf{x}}[\mathbf{M}^{(r-2)}] \right)_{i_{r-1}, i_r} \end{aligned} \quad (3.55)$$

Il s'ensuit que l'ensemble $\mathcal{M}^{(r)}$ des matrices à bloc-diagonaliser est en fait

$$\mathcal{M}^{(r)} = \left\{ \mathcal{Q}_{(r)}^{\mathbf{x}}[\mathbf{E}_{i_1, i_2, \dots, i_{r-2}}] \quad ; \quad 1 \leq i_1 \leq i_2 \leq \dots \leq i_{r-2} \leq n \right\}, \quad (3.56)$$

qui est, en général, de cardinal

$$\text{card}(\mathcal{M}^{(r)}) = \binom{n+r-3}{r-2} = \frac{n(n+1) \dots (n+r-3)}{(r-2)!}, \quad (3.57)$$

qui est aussi le cardinal de l'ensemble des combinaisons avec répétition de n objets pris $r-2$ à $r-2$. Pour plus des détails voir l'annexe (C).

3.7 De l'intérêt de combiner des ordres différents

Puisque les matrices de cumulants contractées de tout ordre sont bloc-diagonalisable dans la même base, rien n'empêche de tenter une résolution d'un problème d'ACIM en bloc-diagonalisant un ensemble

$$\mathcal{M} = \bigcup_{l=1}^q \mathcal{M}^{(r_l)} \quad (3.58)$$

de matrices de cumulants contractées d'ordre r_1, \dots, r_q . Convenons de noter $\text{SJADE}_{r_1, \dots, r_q}$ l'algorithme correspondant. Il s'agit d'une généralisation multidimensionnelle et à des ordres quelconques de la méthode eJADE proposée par [Moreau 2001] pour l'ACI scalaire et qui diagonalise des matrices de cumulants d'ordres 3 et 4.

Restons dans le cadre scalaire pour y présenter une situation où seule une combinaison idoine des cumulants d'ordres variés permet d'envisager la séparation. Considérons $\mathbf{s} = (s_1, s_2, s_3, s_4)'$, où les sources scalaires s_i ($i = 1, \dots, 4$) sont indépendantes et telles que $k_1^{(4)} = k_2^{(4)} = k_3^{(6)} = k_4^{(6)} = 0$ et $k_1^{(6)} \neq 0, k_2^{(6)} \neq 0, k_3^{(4)} \neq 0$ et $k_4^{(4)} \neq 0$ (où $k_i^{(4)} = \text{cum}(s_i, s_i, s_i, s_i)$ et $k_i^{(6)} = \text{cum}(s_i, s_i, s_i, s_i, s_i, s_i)$). De telles sources sont évidemment non gaussiennes. Notons qu'un tel scénario est effectivement possible (voir [Petrochilos 2006]). Posons $\mathbf{x} = \mathbf{A}\mathbf{s}$ où $\mathbf{A} \in \mathbb{R}^{4 \times 4}$. Le corollaire 23 affirme que

$$\mathcal{Q}_{(4)}^{\mathbf{x}}[\mathbf{M}^{(2)}] = \mathbf{A} \text{diag} \left(0, 0, \mathbf{a}_3' \mathbf{M}^{(2)} \mathbf{a}_3 k_3^{(4)}, \mathbf{a}_4' \mathbf{M}^{(2)} \mathbf{a}_4 k_4^{(4)} \right) \mathbf{A}' \quad (3.59)$$

$$\mathcal{Q}_{(6)}^{\mathbf{x}}[\mathbf{M}^{(4)}] = \mathbf{A} \text{diag} \left((\mathbf{a}_1 \otimes \mathbf{a}_1)' \mathbf{M}^{(4)} (\mathbf{a}_1 \otimes \mathbf{a}_1) k_1^{(6)}, (\mathbf{a}_2 \otimes \mathbf{a}_2)' \mathbf{M}^{(4)} (\mathbf{a}_2 \otimes \mathbf{a}_2) k_2^{(6)}, 0, 0 \right) \mathbf{A}' \quad (3.60)$$

Il est clair que la matrice de mélange inconnue \mathbf{A} ne peut pas être estimée dans tel scénario via une méthode d'ACI d'ordre 4 ni d'ordre 6, à cause des deux cumulants nuls dans chaque cas (nécessité de l'hypothèse (h_2) de la section 3.2.1). En revanche, si nous définissons la matrice $\mathcal{Q}_{(4,6)} = \mathcal{Q}_{(4)}^{\mathbf{x}}[\mathbf{M}^{(2)}] + \mathcal{Q}_{(6)}^{\mathbf{x}}[\mathbf{M}^{(4)}]$, alors

$$\mathcal{Q}_{(4,6)} = \mathbf{A} \text{diag} \left((\mathbf{a}_1 \otimes \mathbf{a}_1)' \mathbf{M}^{(4)} (\mathbf{a}_1 \otimes \mathbf{a}_1) k_1^{(6)}, (\mathbf{a}_2 \otimes \mathbf{a}_2)' \mathbf{M}^{(4)} (\mathbf{a}_2 \otimes \mathbf{a}_2) k_2^{(6)}, \mathbf{a}_3' \mathbf{M}^{(2)} \mathbf{a}_3 k_3^{(4)}, \mathbf{a}_4' \mathbf{M}^{(2)} \mathbf{a}_4 k_4^{(4)} \right) \mathbf{A}' \quad (3.61)$$

Dans cette circonstance, la diagonalisation d'un ensemble de matrices de type $\mathcal{Q}_{(4,6)}$ devrait résoudre le problème. Vérifions-le sur un exemple numérique. Pour cela, considérons un vecteur source $\mathbf{s} = (s_1, s_2, s_3, s_4)'$, où les sources scalaires s_i ($i = 1, \dots, 4$) sont indépendantes et telles que $k_1^{(4)} = k_2^{(4)} = k_3^{(6)} = k_4^{(6)} = 0$ et $k_1^{(6)} = 0.9730, k_2^{(6)} = 1.2980, k_3^{(4)} = -0.8003$ et $k_4^{(4)} = 1.3614$. Nous mélangeons ces sources par la matrice suivante :

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 \end{pmatrix} = \begin{pmatrix} 0.9787 & 0.0596 & 0.5216 & 0.7224 \\ 0.7127 & 0.6820 & 0.0967 & 0.1499 \\ 0.5005 & 0.0424 & 0.8181 & 0.6596 \\ 0.4711 & 0.0714 & 0.8175 & 0.5186 \end{pmatrix} \quad (3.62)$$

Ces seules données permettent le calcul des matrices de cumulants d'ordre 4 et 6 du vecteur d'observation $\mathbf{x} = \mathbf{A}\mathbf{s}$:

$$\begin{aligned} c_4(\mathbf{x}) &= \mathcal{A}_2 \text{diag} \left(k_1^{(4)}, k_2^{(4)}, k_3^{(4)}, k_4^{(4)} \right) \mathcal{A}_2' \\ c_6(\mathbf{x}) &= \mathcal{A}_3 \text{diag} \left(k_1^{(6)}, k_2^{(6)}, k_3^{(6)}, k_4^{(6)} \right) \mathcal{A}_3' \end{aligned} \quad (3.63)$$

où

$$\begin{aligned} \mathcal{A}_2 &= \begin{pmatrix} \mathbf{a}_1 \otimes \mathbf{a}_1 & \mathbf{a}_2 \otimes \mathbf{a}_2 & \mathbf{a}_3 \otimes \mathbf{a}_3 & \mathbf{a}_4 \otimes \mathbf{a}_4 \end{pmatrix} \\ \mathcal{A}_3 &= \begin{pmatrix} \mathbf{a}_1 \otimes \mathbf{a}_1 \otimes \mathbf{a}_1 & \mathbf{a}_2 \otimes \mathbf{a}_2 \otimes \mathbf{a}_2 & \mathbf{a}_3 \otimes \mathbf{a}_3 \otimes \mathbf{a}_3 & \mathbf{a}_4 \otimes \mathbf{a}_4 \otimes \mathbf{a}_4 \end{pmatrix} \end{aligned} \quad (3.64)$$

Nous construisons ensuite les deux ensembles de matrices de cumulants contractées $\mathcal{M}^{(4)}$ et $\mathcal{M}^{(6)}$ définis par (3.56); sur lesquels nous utilisons un algorithme de diagonalisation conjointe non orthogonale. Il s'agit ici de l'algorithme *FFDiag* de [Ziehe 2004]. Dans le cas particulier où la matrice \mathbf{A} serait choisie orthogonale, il serait loisible d'utiliser l'algorithme de [Cardoso 1996].

Nous calculons ainsi successivement trois estimations de l'inverse de \mathbf{A} :

- \mathbf{B}_4 , en diagonalisant l'ensemble $\mathcal{M}^{(4)}$;
- \mathbf{B}_6 , en diagonalisant l'ensemble $\mathcal{M}^{(6)}$;
- $\mathbf{B}_{4,6}$, en diagonalisant l'ensemble $\mathcal{M}^{(4,6)} = \mathcal{M}^{(4)} \cup \mathcal{M}^{(6)}$.

Pour tester la qualité des résultats, voici les produits $\mathbf{B}_\bullet \mathbf{A}$ obtenus :

$$\mathbf{B}_4 \mathbf{A} = \begin{pmatrix} -0.5109 & -0.6400 & -0.0000 & 0.0061 \\ 0.5266 & 0.6311 & -0.0000 & 0.0007 \\ 0.0420 & 0.2293 & 0.0001 & 0.0000 \\ -0.0479 & -0.2329 & 0.0039 & 0.0000 \end{pmatrix} \quad (3.65)$$

$$\mathbf{B}_6 \mathbf{A} = \begin{pmatrix} 0.0027 & -0.0000 & -0.0631 & -0.1051 \\ -0.0000 & 0.0000 & -0.3507 & -0.0808 \\ -0.0001 & 0.0000 & 0.6606 & 0.2647 \\ -0.0000 & 0.0043 & 0.3823 & 0.0952 \end{pmatrix} \quad (3.66)$$

$$\mathbf{B}_{4,6} \mathbf{A} = \begin{pmatrix} 0.0000 & -0.1792 & -0.0000 & -0.0000 \\ 0.0000 & 0.0000 & -0.1813 & 0.0000 \\ -0.1435 & -0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & -0.0801 \end{pmatrix} \quad (3.67)$$

Comme attendu, nous constatons que les deux méthodes n'utilisant que l'ordre 4 ou que l'ordre 6 échouent complètement tandis que la méthode qui combine les deux statistiques permet une estimation parfaite (aux indéterminations près de l'ACI) de la matrice de mélange.

3.8 Vers une détermination automatique de la taille des sources ?

Il a été montré dans les références [Theis 2007a, Gutch 2007, Gutch 2012] que l'hypothèse d'irréductibilité de sources joue un rôle clé dans l'ACIM, chaque source \mathbf{s}_i devant être irréductible afin de garantir l'unicité de la solution. Néanmoins, tester l'irréductibilité d'un vecteur aléatoire \mathbf{y} à partir de la définition 3 page 17 est un peu compliqué. Nous allons donner ici un résultat sur le rang de $c_4(\mathbf{y})$, qui garantit l'irréductibilité de \mathbf{y} .

Lemme 26. Soit $\mathbf{y} \in \mathbb{R}^d$ un vecteur aléatoire vérifiant la condition

$$\text{rang}(c_4(\mathbf{y})) = \frac{d(d+1)}{2}. \quad (3.68)$$

Alors \mathbf{y} est irréductible.

Démonstration. Nous allons raisonner par contraposée. En effet, supposons \mathbf{y} réductible, alors d'après la définition 3, il existe une matrice $\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 \end{pmatrix} \in \mathbb{R}^{d \times d}$ inversible et des vecteurs $\mathbf{y}_1 \in \mathbb{R}^{d_1}$ et $\mathbf{y}_2 \in \mathbb{R}^{d_2}$ indépendants ($d = d_1 + d_2$), tels que

$$\mathbf{y} = \mathbf{A} \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix}. \quad (3.69)$$

Ceci implique, d'après (3.12), que

$$c_4(\mathbf{y}) = \underbrace{\begin{pmatrix} \mathbf{A}_1 \otimes \mathbf{A}_1 & \mathbf{A}_2 \otimes \mathbf{A}_2 \end{pmatrix}}_{=\mathcal{A}_2} \text{bdiag} \left(c_4(\mathbf{y}_1), c_4(\mathbf{y}_2) \right) \begin{pmatrix} \mathbf{A}_1 \otimes \mathbf{A}_1 & \mathbf{A}_2 \otimes \mathbf{A}_2 \end{pmatrix}'. \quad (3.70)$$

Par suite,

$$\begin{aligned}
\text{rang}(c_4(\mathbf{y})) &= \text{rang}\left(\mathcal{A}_2 \text{ bdiag}\left(c_4(\mathbf{y}_1), c_4(\mathbf{y}_2)\right) \mathcal{A}'_2\right) \\
&= \text{rang}\left(\text{bdiag}\left(c_4(\mathbf{x}_1), c_4(\mathbf{x}_2)\right)\right) \\
&= \text{rang}(c_4(\mathbf{x}_1)) + \text{rang}(c_4(\mathbf{x}_2)) \\
&\leq \frac{d_1(d_1+1)}{2} + \frac{d_2(d_2+1)}{2} \\
&< \frac{d_1(d_1+1)}{2} + \frac{d_2(d_2+1)}{2} + d_1d_2 = \frac{d(d+1)}{2},
\end{aligned} \tag{3.71}$$

où nous avons utilisé l'égalité $\text{rang}(\mathbf{A}_1 \otimes \mathbf{A}_1 \quad \mathbf{A}_2 \otimes \mathbf{A}_2) = d_1^2 + d_2^2$, valide puisque $\text{rang}(\mathbf{A}) = \text{rang}(\mathbf{A}_1) + \text{rang}(\mathbf{A}_2)$ (voir [Tian 2005]), ainsi que la majoration du rang (voir le lemme 16 page 35). Cela signifie que lorsque \mathbf{y} est réductible, on doit avoir $\text{rang}(c_4(\mathbf{y})) < \frac{d(d+1)}{2}$. Par conséquent, si on a $\text{rang}(c_4(\mathbf{y})) = \frac{d(d+1)}{2}$, alors \mathbf{y} est irréductible. \square

Remarque 27. De la même façon, on peut montrer que si $\mathbf{y} \in \mathbb{R}^d$ est un vecteur aléatoire qui vérifie

$$\text{rang}(c_6(\mathbf{y})) = \frac{d(d+1)(d+2)}{6}, \tag{3.72}$$

alors il est irréductible.

Remarque 28. La condition précédente (3.68) garantit l'irréductibilité d'un vecteur aléatoire (condition suffisante), mais elle n'est pas du tout nécessaire. En effet, rien n'interdit de penser à un vecteur aléatoire qui vérifierait (3.72) -il serait donc irréductible- sans vérifier (3.68). Notons bien qu'en pratique, si on génère « au hasard » un vecteur aléatoire, nous aurons presque sûrement ces deux conditions vérifiées.

Lemme 29. Soit \mathbf{y} un vecteur aléatoire vérifiant (3.68) et soit $\mathbf{U}\mathbf{\Lambda}\mathbf{U}'$ une décomposition en valeurs propres de $c_4(\mathbf{y})$, où $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_{\frac{d(d+1)}{2}})$, alors nous avons

$$\mathbf{U}\mathbf{U}' = \frac{1}{2}(\mathbf{K}_{\text{dd}} + \mathbf{I}_{d^2}), \tag{3.73}$$

où \mathbf{K}_{dd} est une matrice de commutation définie par la formule (2.20) page 27.

Démonstration. Comme $c_4(\mathbf{y})$ est une matrice symétrique de rang $\frac{d(d+1)}{2}$, elle admet une décomposition en valeurs propres $\mathbf{U}\mathbf{\Lambda}\mathbf{U}'$, où $\mathbf{U} \in \mathbb{R}^{d^2 \times \frac{d(d+1)}{2}}$ est une matrice semi-orthogonale ($\mathbf{U}'\mathbf{U} = \mathbf{I}_{\frac{d(d+1)}{2}}$) et $\mathbf{\Lambda} \in \mathbb{R}^{\frac{d(d+1)}{2} \times \frac{d(d+1)}{2}}$ est une matrice diagonale constituée des $\frac{d(d+1)}{2}$ valeurs propres non nulles de $c_4(\mathbf{y})$. Soit $\tilde{\mathbf{K}} = \frac{1}{2}(\mathbf{K}_{\text{dd}} + \mathbf{I}_{d^2})$. Comme $\tilde{\mathbf{K}}c_4(\mathbf{y})\tilde{\mathbf{K}} = c_4(\mathbf{y})$, on a $\tilde{\mathbf{K}}\mathbf{U}\mathbf{\Lambda}\mathbf{U}'\tilde{\mathbf{K}} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}'$. Ceci entraîne que $\mathbf{U}'\tilde{\mathbf{K}}\tilde{\mathbf{K}}\mathbf{U} = \mathbf{I}_{\frac{d(d+1)}{2}}$. Comme $\tilde{\mathbf{K}}$ est idempotente ($\tilde{\mathbf{K}}\tilde{\mathbf{K}} = \tilde{\mathbf{K}}$), il vient $\mathbf{U}'\tilde{\mathbf{K}}\mathbf{U} = \mathbf{I}_{\frac{d(d+1)}{2}}$, et donc $\mathbf{U}\mathbf{U}'\tilde{\mathbf{K}}\mathbf{U}\mathbf{U}' = \mathbf{U}\mathbf{U}'$. Ainsi $\mathbf{U}\mathbf{U}'$ est le pseudo-inverse de $\tilde{\mathbf{K}}$ qui est égal à $\tilde{\mathbf{K}}$. \square

Remarque 30. La matrice $\tilde{\mathbf{K}} = \frac{1}{2}(\mathbf{K}_{\text{dd}} + \mathbf{I}_{d^2})$ est appelée en anglais *the symmetrizer matrix* (pour plus de détails, voir par exemple [Abadir 2005]) et souvent notée \mathbf{N}_d .

Supposons maintenant que toutes les sources indépendantes (multidimensionnelles) vérifient la condition (3.68), c'est-à-dire $\text{rang}(c_4(\mathbf{s}_i)) = \frac{n_i(n_i+1)}{2}$; elles sont donc irréductibles. On sait alors automatiquement qu'aucune source n'est gaussienne (puisque $\text{rang}(c_4(\mathbf{s}_i)) \neq 0$). De plus, à partir du modèle ACIM blanchie ($\mathbf{x} = \mathbf{A}\mathbf{s}$), nous avons d'après (3.27) l'expression de $c_4(\mathbf{x})$:

$$c_4(\mathbf{x}) = \begin{pmatrix} \mathbf{A}\mathbf{D}_1 \\ \vdots \\ \mathbf{A}\mathbf{D}_n \end{pmatrix} \begin{pmatrix} c_4(\mathbf{s}_1) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & c_4(\mathbf{s}_p) \end{pmatrix} \begin{pmatrix} \mathbf{D}'_1\mathbf{A}' & \cdots & \mathbf{D}'_n\mathbf{A}' \end{pmatrix} \quad (3.74)$$

Désignons par $\mathbf{U}_i\Lambda_i\mathbf{U}'_i$, pour $i = 1, \dots, p$, des décompositions en valeurs propres des $c_4(\mathbf{s}_i)$, comme dans le lemme 29 : $\mathbf{U}_i\mathbf{U}'_i = \frac{1}{2}(\mathbf{K}_{n_i n_i} + \mathbf{I}_{n_i^2})$. Posons

$$\mathbf{U} = \text{bdiag}(\mathbf{U}_1, \dots, \mathbf{U}_p) \quad (3.75)$$

et

$$\Lambda = \text{bdiag}(\Lambda_1, \dots, \Lambda_p), \quad (3.76)$$

de sorte que l'équation (3.74) devienne

$$c_4(\mathbf{x}) = \begin{pmatrix} \mathbf{A}\mathbf{D}_1\mathbf{U} \\ \vdots \\ \mathbf{A}\mathbf{D}_n\mathbf{U} \end{pmatrix} \begin{pmatrix} \Lambda_1 & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \Lambda_p \end{pmatrix} \begin{pmatrix} \mathbf{U}'\mathbf{D}'_1\mathbf{A}' & \cdots & \mathbf{U}'\mathbf{D}'_n\mathbf{A}' \end{pmatrix}. \quad (3.77)$$

On reconnait une décomposition en valeurs propres de $c_4(\mathbf{x}) = \mathbf{V}\Lambda\mathbf{V}'$, en posant

$$\mathbf{V} = \begin{pmatrix} \mathbf{A}\mathbf{D}_1\mathbf{U} \\ \mathbf{A}\mathbf{D}_2\mathbf{U} \\ \vdots \\ \mathbf{A}\mathbf{D}_n\mathbf{U} \end{pmatrix}. \quad (3.78)$$

Calculons le produit

$$\mathbf{V}'\mathbf{V} = \sum_{k=1}^n \mathbf{U}'\mathbf{D}'_k \underbrace{\mathbf{A}'\mathbf{A}}_{=\mathbf{I}_n} \mathbf{D}_k\mathbf{U} = \mathbf{U}' \left(\sum_{k=1}^n \mathbf{D}'_k\mathbf{D}_k \right) \mathbf{U}, \quad (3.79)$$

En remplaçant maintenant \mathbf{D}_k par sa valeur (3.30), nous obtenons

$$\begin{aligned} \mathbf{V}'\mathbf{V} &= \mathbf{U}' \left(\sum_{k=1}^n \text{bdiag}(\ell_1^k \otimes \mathbf{I}_{n_1}, \dots, \ell_p^k \otimes \mathbf{I}_{n_p})' \text{bdiag}(\ell_1^k \otimes \mathbf{I}_{n_1}, \dots, \ell_p^k \otimes \mathbf{I}_{n_p}) \right) \mathbf{U} \\ &= \mathbf{U}' \left(\sum_{k=1}^n \text{bdiag}(\ell_1^{k'} \ell_1^k, \dots, \ell_p^{k'} \ell_p^k) \right) \mathbf{U} \\ &= \mathbf{U}' \text{bdiag} \left(\sum_{k=1}^n \ell_1^{k'} \ell_1^k, \dots, \sum_{k=1}^n \ell_p^{k'} \ell_p^k \right) \mathbf{U} \\ &= \mathbf{U}' \text{bdiag}(\mathbf{A}'_1\mathbf{A}_1, \dots, \mathbf{A}'_p\mathbf{A}_p) \mathbf{U} \\ &= \mathbf{U}' \text{bdiag}(\mathbf{I}_{n_1}, \dots, \mathbf{I}_{n_p}) \mathbf{U} \\ &= \mathbf{U}'\mathbf{U} \\ &= \text{bdiag}(\mathbf{U}'_1\mathbf{U}_1, \dots, \mathbf{U}'_p\mathbf{U}_p) \\ &= \text{bdiag}(\mathbf{I}_{\frac{n_1(n_1+1)}{2}}, \dots, \mathbf{I}_{\frac{n_p(n_p+1)}{2}}) \\ &= \mathbf{I}_{\sum_{i=1}^p \frac{n_i(n_i+1)}{2}} \end{aligned} \quad (3.80)$$

La matrice $\mathbf{V}'\mathbf{V}$ est partagée en n^2 matrices bloc de taille $n \times n$ données par

$$[\mathbf{V}'\mathbf{V}]_{i,j} = \mathbf{A}\mathbf{D}_i\mathbf{U}\mathbf{U}'\mathbf{D}'_j\mathbf{A}' \quad (3.81)$$

D'un autre côté, d'après le lemme 29, on a

$$\mathbf{U}\mathbf{U}' = \text{bdiag}(\mathbf{U}_1\mathbf{U}'_1, \dots, \mathbf{U}_p\mathbf{U}'_p) = \text{bdiag}\left(\frac{1}{2}(\mathbf{K}_{n_1 n_1} + \mathbf{I}_{n_1^2}), \dots, \frac{1}{2}(\mathbf{K}_{n_p n_p} + \mathbf{I}_{n_p^2})\right) \quad (3.82)$$

Par suite,

$$\left[\mathbf{V}'\mathbf{V}\right]_{i,j} = \mathbf{A} \text{ bdiag} \left((\ell_1^i \otimes \mathbf{I}_{n_1}) \frac{1}{2}(\mathbf{K}_{n_1 n_1} + \mathbf{I}_{n_1^2}) (\ell_1^j \otimes \mathbf{I}_{n_1})', \dots, (\ell_p^i \otimes \mathbf{I}_{n_p}) \frac{1}{2}(\mathbf{K}_{n_p n_p} + \mathbf{I}_{n_p^2}) (\ell_p^j \otimes \mathbf{I}_{n_p})' \right) \mathbf{A}', \quad (3.83)$$

et en utilisant la propriété (P4), du produit étoile (voir (B.3)), nous obtenons

$$\left[\mathbf{V}'\mathbf{V}\right]_{i,j} = \mathbf{A} \text{ bdiag} \left(\ell_1^{i'} \ell_1^j * \frac{1}{2}(\mathbf{K}_{n_1 n_1} + \mathbf{I}_{n_1^2}), \dots, \ell_p^{i'} \ell_p^j * \frac{1}{2}(\mathbf{K}_{n_p n_p} + \mathbf{I}_{n_p^2}) \right) \mathbf{A}' \quad (3.84)$$

Par suite,

$$\left[\mathbf{V}'\mathbf{V}\right]_{i,j} = \mathbf{A} \text{ bdiag} \left(\frac{1}{2}(\ell_1^{j'} \ell_1^i + \ell_1^j \ell_1^{i'} \mathbf{I}_{n_1}), \dots, \frac{1}{2}(\ell_p^{j'} \ell_p^i + \ell_p^j \ell_p^{i'} \mathbf{I}_{n_p}) \right) \mathbf{A}'. \quad (3.85)$$

Proposition 31. Soit \mathcal{R} la matrice définie par

$$\mathcal{R} = \sum_{k=1}^n \left[\mathbf{V}'\mathbf{V}\right]_{k,k}. \quad (3.86)$$

Alors, nous avons

$$\mathcal{R} = \mathbf{A} \begin{pmatrix} \frac{n_1+1}{2} \mathbf{I}_{n_1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \frac{n_2+1}{2} \mathbf{I}_{n_2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \frac{n_p+1}{2} \mathbf{I}_{n_p} \end{pmatrix} \mathbf{A}'. \quad (3.87)$$

Démonstration. D'après (3.84), nous avons

$$\begin{aligned} \mathcal{R} &= \sum_{k=1}^n \left[\mathbf{V}'\mathbf{V}\right]_{k,k} = \sum_{k=1}^n \mathbf{A} \text{ bdiag} \left(\ell_1^k \ell_1^k * \frac{1}{2}(\mathbf{K}_{n_1 n_1} + \mathbf{I}_{n_1^2}), \dots, \ell_p^k \ell_p^k * \frac{1}{2}(\mathbf{K}_{n_p n_p} + \mathbf{I}_{n_p^2}) \right) \mathbf{A}' \\ &= \mathbf{A} \text{ bdiag} \left(\underbrace{\sum_{k=1}^n \ell_1^k \ell_1^k * \frac{1}{2}(\mathbf{K}_{n_1 n_1} + \mathbf{I}_{n_1^2})}_{=\mathbf{A}'_1 \mathbf{A}_1}, \dots, \underbrace{\sum_{k=1}^n \ell_p^k \ell_p^k * \frac{1}{2}(\mathbf{K}_{n_p n_p} + \mathbf{I}_{n_p^2})}_{=\mathbf{A}'_p \mathbf{A}_p} \right) \mathbf{A}' \\ &= \mathbf{A} \text{ bdiag} \left((\mathbf{A}'_1 \mathbf{A}_1) * \frac{1}{2}(\mathbf{K}_{n_1 n_1} + \mathbf{I}_{n_1^2}), \dots, (\mathbf{A}'_p \mathbf{A}_p) * \frac{1}{2}(\mathbf{K}_{n_p n_p} + \mathbf{I}_{n_p^2}) \right) \mathbf{A}' \\ &= \mathbf{A} \text{ bdiag} \left(\mathbf{I}_{n_1} * \frac{1}{2}(\mathbf{K}_{n_1 n_1} + \mathbf{I}_{n_1^2}), \dots, \mathbf{I}_{n_p} * \frac{1}{2}(\mathbf{K}_{n_p n_p} + \mathbf{I}_{n_p^2}) \right) \mathbf{A}' \\ &= \mathbf{A} \text{ bdiag} \left(\frac{n_1+1}{2} \mathbf{I}_{n_1}, \dots, \frac{n_p+1}{2} \mathbf{I}_{n_p} \right) \mathbf{A}' \end{aligned} \quad (3.88)$$

□

La proposition 31 signifie que la matrice orthogonale recherchée \mathbf{A} , qui bloc-diagonalise conjointement l'ensemble $\mathcal{M}^{(4)}$, diagonalise aussi la matrice \mathcal{R} . En effet, l'équation (3.87) représente une décomposition en valeurs propres de la matrice symétrique \mathcal{R} ; elle montre aussi que \mathcal{R} admet uniquement p valeurs propres positives, $\frac{n_1+1}{2}, \dots, \frac{n_p+1}{2}$, de multiplicités respectives n_1, \dots, n_p . Le spectre de \mathcal{R} sera noté :

$$\lambda_{\mathcal{R}} = \left\{ \frac{n_1+1}{2}, \dots, \frac{n_p+1}{2} \right\}. \quad (3.89)$$

Remarque 32. On peut remarquer que dans le cas ACI, nous aurons $\mathcal{R} = \mathbf{I}_n$, tandis que dans le cadre k -ACIM nous obtenons $\mathcal{R} = \frac{k+1}{2} \mathbf{I}_n$. Dans ces deux cas, la matrice \mathcal{R} ne permet pas l'identification de la permutation des sous-espaces indépendants. En revanche, si toutes les sources ont des dimensions deux à deux différentes ou égales à 1, si \mathbf{E}' est la solution donnée par l'algorithme de diagonalisation conjointe de type Jacobi [Cardoso 1998], alors $\mathbf{P}'\mathbf{E}'$ est la solution du problème de BDC où \mathbf{P} est la permutation qui ordonne les éléments diagonaux de $\mathbf{E}'\mathcal{R}\mathbf{E}$, c'est-à-dire les valeurs propres de \mathcal{R} (voir la conjecture 11 de [Abed-Meraim 2004]).

3.8.1 Un exemple numérique

Considérons le mélange

$$\mathbf{x} = \mathbf{A} \begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \mathbf{s}_3 \\ \mathbf{s}_4 \end{pmatrix}, \quad (3.90)$$

où les sources \mathbf{s}_i , $i = 1, \dots, 4$ sont indépendantes, vérifient (3.68), elles sont donc irréductibles, de dimensions $\dim(\mathbf{s}_1) = \dim(\mathbf{s}_2) = 1$, $\dim(\mathbf{s}_3) = 2$, $\dim(\mathbf{s}_4) = 3$, et de matrices de cumulants

$$c_4(\mathbf{s}_1) = 1.2331; \quad c_4(\mathbf{s}_2) = -0.5013; \quad c_4(\mathbf{s}_3) = \begin{pmatrix} 0.1449 & 0.5279 & 0.5279 & 0.2170 \\ 0.5279 & 0.2170 & 0.2170 & 0.6261 \\ 0.5279 & 0.2170 & 0.2170 & 0.6261 \\ 0.2170 & 0.6261 & 0.6261 & 0.4908 \end{pmatrix};$$

$$c_4(\mathbf{s}_4) = \begin{pmatrix} 0.9636 & 0.2654 & 0.5322 & 0.2654 & 0.3258 & 0.5285 & 0.5322 & 0.5285 & 0.5011 \\ 0.2654 & 0.3258 & 0.5285 & 0.3258 & 0.6258 & 0.5011 & 0.5285 & 0.5011 & 0.5933 \\ 0.5322 & 0.5285 & 0.5011 & 0.5285 & 0.5011 & 0.5933 & 0.5011 & 0.5933 & 0.5717 \\ 0.2654 & 0.3258 & 0.5285 & 0.3258 & 0.6258 & 0.5011 & 0.5285 & 0.5011 & 0.5933 \\ 0.3258 & 0.6258 & 0.5011 & 0.6258 & 0.7213 & 0.3073 & 0.5011 & 0.3073 & 0.8174 \\ 0.5285 & 0.5011 & 0.5933 & 0.5011 & 0.3073 & 0.8174 & 0.5933 & 0.8174 & 0.2943 \\ 0.5322 & 0.5285 & 0.5011 & 0.5285 & 0.5011 & 0.5933 & 0.5011 & 0.5933 & 0.5717 \\ 0.5285 & 0.5011 & 0.5933 & 0.5011 & 0.3073 & 0.8174 & 0.5933 & 0.8174 & 0.2943 \\ 0.5011 & 0.5933 & 0.5717 & 0.5933 & 0.8174 & 0.2943 & 0.5717 & 0.2943 & 0.4781 \end{pmatrix}.$$

La matrice de mélange \mathbf{A} est une matrice 7×7 orthogonale dont l'expression numérique est :

$$\mathbf{A} = \begin{pmatrix} -0.4658 & 0.1112 & -0.558 & -0.1817 & 0.2191 & 0.4352 & 0.4346 \\ -0.4135 & -0.5279 & 0.1581 & -0.3195 & -0.1066 & 0.3515 & -0.5369 \\ -0.3009 & 0.341 & 0.3417 & -0.317 & 0.6814 & -0.3005 & -0.1464 \\ -0.3266 & 0.1716 & -0.5732 & 0.2306 & -0.146 & -0.4625 & -0.4969 \\ -0.4304 & -0.4736 & 0.1607 & 0.0361 & -0.1486 & -0.5413 & 0.4982 \\ -0.2615 & -0.0762 & 0.2171 & 0.8426 & 0.3111 & 0.2634 & -0.0504 \\ -0.4014 & 0.5772 & 0.3813 & 0.0024 & -0.5797 & 0.1432 & 0.0614 \end{pmatrix}.$$

Ces seules données permettent le calcul de la matrice de cumulants $c_4(\mathbf{x}) = \mathcal{A}_2 \text{bdiag}(c_4(\mathbf{s}_1), \dots, c_4(\mathbf{s}_4)) \mathcal{A}'_2$ (voir (3.27) page 46), ainsi que celui de l'ensemble des matrices de cumulants contractées $\mathcal{M}^{(4)}$. L'algorithme de diagonalisation conjointe [Cardoso 1993], appliqué à l'ensemble $\mathcal{M}^{(4)}$, produit la matrice suivante

$$\mathbf{E} = \begin{pmatrix} 0.5301 & -0.4458 & 0.2827 & -0.4352 & 0.4346 & 0.0333 & 0.247 \\ 0.4377 & 0.5072 & 0.1577 & -0.3515 & -0.5369 & -0.2976 & 0.161 \\ -0.2347 & 0.0354 & 0.6996 & 0.3005 & -0.1464 & 0.2745 & 0.5169 \\ 0.4274 & -0.5085 & -0.2677 & 0.4625 & -0.4969 & 0.0534 & 0.1533 \\ 0.4142 & 0.4731 & -0.1286 & 0.5413 & 0.4982 & -0.0827 & 0.2001 \\ 0.0582 & 0.2287 & -0.3952 & -0.2634 & -0.0504 & 0.8066 & 0.2562 \\ -0.3386 & -0.0907 & -0.4017 & -0.1432 & 0.0614 & -0.418 & 0.7188 \end{pmatrix}.$$

Posons $r = \text{rang}(c_4(\mathbf{x})) = \text{rang}(c_4(\mathbf{s}_1)) + \text{rang}(c_4(\mathbf{s}_2)) + \text{rang}(c_4(\mathbf{s}_3)) + \text{rang}(c_4(\mathbf{s}_4)) = 1 + 1 + 3 + 6 = 11$ et calculons $\mathbf{V}\mathbf{\Lambda}\mathbf{V}'$ une décomposition en $r = 11$ valeurs propres de $c_4(\mathbf{x})$ (avec Matlab, par exemple, il suffit d'utiliser la commande $[\mathbf{V}, \mathbf{\Lambda}] = \text{eigs}(c_4(\mathbf{x}), r)$). On peut alors calculer la matrice \mathcal{R} à partir de (3.86) ainsi que $\mathbf{E}'\mathcal{R}\mathbf{E}$; nous obtenons :

$$\mathcal{R} = \begin{pmatrix} 1.5813 & 0.0631 & 0.0909 & 0.4541 & 0.0386 & -0.0503 & -0.0253 \\ 0.0631 & 1.5314 & 0.0127 & -0.0752 & 0.4556 & 0.0315 & -0.0479 \\ 0.0909 & 0.0127 & 1.6059 & -0.1254 & -0.0334 & 0.0993 & 0.2499 \\ 0.4541 & -0.0752 & -0.1254 & 1.502 & -0.0178 & 0.0223 & 0.0542 \\ 0.0386 & 0.4556 & -0.0334 & -0.0178 & 1.4471 & 0.1756 & 0.0038 \\ -0.0503 & 0.0315 & 0.0993 & 0.0223 & 0.1756 & 1.5247 & 0.0546 \\ -0.0253 & -0.0479 & 0.2499 & 0.0542 & 0.0038 & 0.0546 & 1.8077 \end{pmatrix}$$

$$\text{et} \quad \mathbf{E}'\mathcal{R}\mathbf{E} = \begin{pmatrix} 2 & -0 & -0 & 0 & -0 & 0 & 0 \\ -0 & 2 & 0 & -0 & 0 & -0 & -0 \\ -0 & 0 & 1.5 & -0 & 0 & -0 & 0 \\ 0 & -0 & -0 & 1 & -0 & 0 & -0 \\ -0 & 0 & 0 & -0 & 1 & -0 & 0 \\ 0 & -0 & -0 & 0 & -0 & 1.5 & -0 \\ 0 & -0 & 0 & -0 & 0 & -0 & 2 \end{pmatrix}.$$

La permutation qui ordonne les valeurs propres de \mathcal{R} est

$$\mathbf{P} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Une simple vérification montre alors que

$$\mathbf{P}'\mathbf{E}'\mathbf{A} = \begin{pmatrix} -0.8289 & 0.4456 & 0.3382 & 0 & 0 & -0 & -0 \\ -0.5545 & -0.5748 & -0.6017 & -0 & -0 & 0 & 0 \\ -0.0737 & -0.6863 & 0.7236 & 0 & -0 & -0 & -0 \\ 0 & 0 & 0 & -0.7239 & 0.6899 & -0 & -0 \\ -0 & -0 & -0 & 0.6899 & 0.7239 & 0 & 0 \\ 0 & 0 & 0 & -0 & -0 & -0 & 1 \\ -0 & 0 & -0 & 0 & 0 & -1 & -0 \end{pmatrix},$$

ce qui signifie que la matrice $\mathbf{B} = \mathbf{P}'\mathbf{E}'$ fournit bien une solution du problème ACIM aux indéterminations près.

3.8.2 Estimation sans la condition de rang de $c_4(\mathbf{x})$

L'exemple précédent a été élaboré de façon à satisfaire exactement la condition d'irréductibilité des sources (3.68) page 53 ($\text{rang}(c_4(\mathbf{x})) = \sum \frac{n_i(n_i+1)}{2}$); ceci est évidemment impossible à réaliser sur des données réelles. Pour se placer dans une situation plus réaliste, nous ajoutons à la matrice $c_4(\mathbf{x})$ une perturbation additionnelle de la forme $\varepsilon c_4(\mathbf{z})$, où \mathbf{z} modélise un vecteur bruit indépendant du vecteur des sources \mathbf{s} et ε est un (petit) paramètre positif.

$$c_4 = c_4(\mathbf{x}) + \varepsilon c_4(\mathbf{z}) \quad (3.91)$$

Dans cette situation, on aura en général $\text{rang}(c_4) = \frac{n(n+1)}{2}$. Après calcul de la matrice $\hat{\mathbf{E}}$ par l'algorithme de digonalisation conjointe [Cardoso 1993], il faut estimer r le rang de $c_4(\mathbf{x})$. L'idée la plus simple est de faire varier r de $r = n$ qui est le cas où il y a n sources mutuellement indépendantes (ACI), jusqu'à $r = \frac{n(n-1)}{2} + 1$, qui est le cas où nous avons uniquement deux sources indépendantes dans le mélange de dimensions respectivement $(n-1)$ et 1. Nous calculons dans chaque cas la décomposition en valeurs propres $[\mathbf{V}_r, \mathbf{\Lambda}_r] = \text{eigs}(c_4, r)$, puis la matrice \mathcal{R}_r correspondante (voir (3.86)). D'une autre part, comme nous savons que la matrice $\hat{\mathbf{E}}' \mathcal{R}_r \hat{\mathbf{E}}$ est diagonale pour la vraie valeur de r (voir la proposition 31), il suffit de choisir la valeur de r pour laquelle cette matrice est « au plus proche » d'une matrice diagonale. Nous testons cette proximité par l'indice d'Amari [Amari 1996]; rappelons que celui-ci est compris entre 0 et 1, d'autant plus proche de 0 que la matrice est diagonale.

Voici un exemple pour lequel $\mathbf{z} = \text{randn}(7, 1000)$ et $\varepsilon = 10^{-2}$; le rang de la matrice de cumulants des observations est $\text{rang}(c_4) = \text{rang}(c_4(\mathbf{x}) + \varepsilon c_4(\mathbf{z})) = \frac{n(n+1)}{2} = 28$. En variant r de n à $\frac{n(n-1)}{2} + 1$ (ici $n = 7$), voici les indices d'Amari observés :

$$\{0.038, 0.048, 0.022, 0.016, 0.001, 0.025, 0.035, 0.035, 0.040, 0.044, 0.048, 0.046, 0.040, 0.039, 0.033, 0.032\}.$$

Le minimum observé (ici 0.001) se trouve dans la case numéro 4 (la première case porte le numéro 0 et correspond à $r = n = 7$); ce qui veut dire que $r = n + 4 = 11$. Nous calculons alors la matrice

$$\hat{\mathbf{E}}' \mathcal{R}_{11} \hat{\mathbf{E}} = \begin{pmatrix} \mathbf{2.0000} & 0.0000 & 0.0098 & -0.0027 & 0.0015 & -0.0045 & -0.0001 \\ 0.0000 & \mathbf{2.0000} & 0.0054 & -0.0016 & -0.0003 & 0.0016 & -0.0000 \\ 0.0098 & 0.0054 & \mathbf{1.4996} & -0.0012 & 0.002 & 0.0003 & -0.0106 \\ -0.0027 & -0.0016 & -0.0012 & \mathbf{1.0002} & -0.0013 & -0.0015 & -0.0010 \\ 0.0015 & -0.0003 & 0.002 & -0.0013 & \mathbf{1.0002} & -0.0004 & 0.0029 \\ -0.0045 & 0.0016 & 0.0003 & -0.0015 & -0.0004 & \mathbf{1.4998} & 0.0096 \\ -0.0001 & -0.0000 & -0.0106 & -0.001 & 0.0029 & 0.0096 & \mathbf{2.0002} \end{pmatrix},$$

effectivement assez proche d'une matrice diagonale, dont on ordonne les éléments diagonaux par ordre décroissant via la permutation \mathbf{P} . La matrice de séparation estimée est ici

$$\hat{\mathbf{B}} = \mathbf{P}' \hat{\mathbf{E}}' = \begin{pmatrix} 0.5301 & 0.4373 & -0.2347 & 0.4274 & 0.4145 & 0.0581 & -0.3388 \\ -0.4454 & 0.5069 & 0.0356 & -0.5085 & 0.4736 & 0.2288 & -0.0902 \\ 0.2472 & 0.1609 & 0.5172 & 0.1539 & 0.1997 & 0.2564 & 0.7185 \\ 0.033 & -0.2975 & 0.2746 & 0.0534 & -0.083 & 0.8063 & -0.4185 \\ 0.2815 & 0.1595 & 0.6997 & -0.2666 & -0.1303 & -0.3953 & -0.4019 \\ 0.4357 & -0.5368 & -0.144 & -0.4975 & 0.4974 & -0.0512 & 0.0606 \\ -0.4351 & -0.3518 & 0.3008 & 0.4622 & 0.5412 & -0.2637 & -0.1433 \end{pmatrix};$$

nous vérifions sa qualité par le calcul du produit

$$\hat{\mathbf{B}}\mathbf{A} = \begin{pmatrix} -0.5544 & -0.5749 & -0.6018 & 0.0000 & 0.0002 & -0.0003 & 0.0003 \\ -0.0742 & -0.6861 & 0.7237 & 0.0001 & -0.0001 & -0.0001 & 0.0006 \\ -0.829 & 0.4459 & 0.3377 & 0.0001 & 0.0005 & -0.0001 & -0.0005 \\ 0.0005 & -0.0002 & -0.0001 & 0.6897 & 0.7241 & -0.0001 & -0.0004 \\ 0.0003 & -0.0001 & 0.0000 & -0.7241 & 0.6897 & 0.0005 & -0.0029 \\ -0.0002 & 0.0008 & -0.0000 & -0.0018 & 0.0023 & 0.0003 & 1.0000 \\ 0.0003 & 0.0002 & 0.0001 & -0.0004 & 0.0003 & -1.0000 & 0.0003 \end{pmatrix}.$$

3.9 Etude expérimentale

3.9.1 Indice de performance pour l'ACIM

Pour mesurer la qualité de la séparation dans les simulations informatiques, nous allons généraliser l'indice de performance proposé par [Moreau 2001] pour l'ACI ordinaire. Autrement dit nous mesurons la proximité de $\mathbf{G} = \mathbf{B}\mathbf{A}$, où \mathbf{B} est la matrice de séparation, à un produit $\mathbf{P}\mathbf{D}$ d'une matrice de permutation \mathbf{P} par une matrice bloc-diagonale \mathbf{D} . Une bonne séparation est obtenue lorsque la matrice \mathbf{G} est égale à une matrice contenant un seul bloc non-nul par bloc-ligne et par bloc-colonne. L'idée que nous allons exploiter ici a été introduite initialement par [Zoltan Szabo 2012]. Formellement, supposons sans perte de généralité que les dimensions des sources (multidimensionnelles) et leurs estimations sont classées dans l'ordre croissant $n_1 \leq n_2 \leq \dots \leq n_p$. Nous pouvons considérer la matrice \mathbf{G} , de taille $n \times n$, comme une matrice partagée en blocs \mathbf{G}^{ij} de taille $n_i \times n_j$ avec $(i, j) \in \llbracket 1, p \rrbracket^2$. Définissons enfin une matrice \mathcal{G} , de taille $p \times p$, par

$$\mathcal{G}_{ij} = \frac{1}{n_i n_j} \sum_{k,l=1}^{n_i, n_j} |(\mathbf{G}^{ij})_{kl}| \quad (3.92)$$

Dire que la matrice \mathbf{G} est bloc-diagonale (pour les tailles de bloc $n_1 \leq n_2 \leq \dots \leq n_p$) équivaut à dire que la matrice \mathcal{G} est diagonale. Ainsi, nous définirons l'indice de performance $I(\mathbf{G})$ pour l'ACIM par

$$I(\mathbf{G}) = \frac{1}{2p(p-1)} \left[\sum_i^p \left(\sum_j^p \frac{\mathcal{G}_{ij}}{\max_l \mathcal{G}_{il}} - 1 \right) + \sum_j^p \left(\sum_i^p \frac{\mathcal{G}_{ij}}{\max_l \mathcal{G}_{lj}} - 1 \right) \right] \quad (3.93)$$

Une séparation parfaite correspond à un indice de performance tendant vers 0 (et donc $-\infty$ en échelle logarithmique). Notons que lorsque $\forall i, n_i = 1$, nous retrouvons l'indice de performance introduit dans [Moreau 2001], utilisé pour l'ACI scalaire.

3.9.2 Comparaisons dans le cadre scalaire (ACI)

Nous allons maintenant comparer les performances des méthodes que nous venons de présenter (on les désignera par JADE_k au lieu de SJADE_k quand il s'agit du cas scalaire) entre elles, puis avec les méthodes classiques : SOBI [A. Belouchrani 1997], JADE [Cardoso 1993], JADEop, JADETD [Georgiev 2003], ERICA [Ding 2006], SIMBEC [S. Cruces 2001], FastICA [Hyvarinen 1999] et Fixed-Point ICA [Hyvärinen 1997].

3.9.2.1 Première simulation

Nous considérons ces trois sources synthétiques scalaires,

$$\begin{aligned} s_1 &= \exp(2(0.007t + 0.5 - \text{floor}(0.007t + 0.5)) - 1) \\ s_2 &= \exp(\cos(0.3t)) \\ s_3 &= \text{square}(0.2t), \end{aligned} \quad (3.94)$$

échantillonnées sur 3000 points (figure 3.1 (a)). Ces trois sources ont été mélangées (figure 3.1 (b)) par une matrice arbitraire $\mathbf{A} : 3 \times 3$, générée avec Matlab via la commande $\mathbf{A} = \text{randn}(3)$.

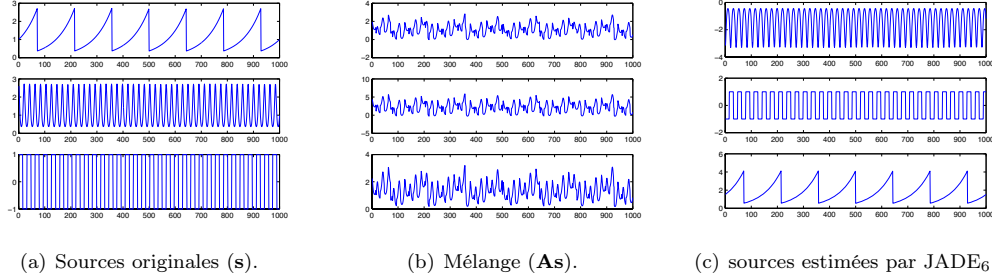


FIGURE 3.1 – *ACI, première simulation.*

Nous avons alors appliqué les méthodes SOBI, JADE₃, JADE, JADE₅ et JADE₆ sur ce mélange. L'indice de performance $I(\mathbf{G})$ de chacune de ces méthodes est donné dans le tableau 3.1. On observe un très léger avantage des méthodes d'ordre pair SOBI et JADE par rapport aux méthodes exploitant des ordres impairs JADE₃ et JADE₅. Cependant la meilleure performance a été enregistrée par la méthode JADE₆ (figure 3.1 (c)).

| Méthode | SOBI | JADE ₃ | JADE | JADE ₅ | JADE ₆ |
|---------------------|--------|-------------------|---------|-------------------|-------------------|
| $I(\mathbf{G})[dB]$ | -45.32 | -43.40 | - 45.94 | -41.31 | -52.31 |

TABLE 3.1 – *Performances (ACI, première simulation)*

3.9.2.2 Seconde simulation

Nous savons que les méthodes d'ordre impair échouent complètement à séparer un mélange qui contient plus d'une source à distribution symétrique. Par contre, les méthodes d'ordre pair ne souffrent pas de ce problème ; le seul type de source à distribution symétrique à éviter sont les gaussiennes. Ceci donne un avantage considérable aux méthodes d'ordre pair par rapport à celles d'ordre impair. Nous testons cela par cette deuxième simulation où nous avons mélangé trois sources scalaires : une gaussienne ($s_1 = \text{randn}(1, 1000)$) avec deux uniformes indépendantes ($s_i = \text{rand}(1, 1000)$, $i = 2, 3$). Les critères de performance des différentes méthodes de séparation sont regroupés dans le tableau 3.2. Comme attendu, nous observons la supériorité écrasante des méthodes JADE et JADE₆ par rapport aux méthodes SOBI, JADE₃ et JADE₅. Remarquons que la méthode SOBI a échoué à cause de la non existence d'une diversité temporelle (les sources ne sont pas colorées et de spectres distincts).

| Méthode | SOBI | JADE ₃ | JADE | JADE ₅ | JADE ₆ |
|---------------------|-------|-------------------|---------|-------------------|-------------------|
| $I(\mathbf{G})[dB]$ | -4.53 | -5.66 | - 26.93 | -4.97 | -26.91 |

TABLE 3.2 – *Performances (ACI, seconde simulation)*

3.9.2.3 Troisième simulation

Néanmoins, il existe des situations où des méthodes d'ordre impair s'avèrent plus performantes. Afin d'éclairer ce point, nous allons faire une analyse Monte Carlo en utilisant la boîte à outils ICALAB [Cichocki 2007]. Il s'agit d'un package Matlab qui permet de comparer les algorithmes de séparation de sources. La qualité de l'extraction d'une source est mesurée en utilisant le Rapport Signal à Interférence (SIR en anglais), pour la matrice de mélange estimée. Des valeurs plus élevées de SIR indiquent une meilleure performance. Plus précisément, une bonne séparation est obtenue si on arrive à estimer le mélange avec un $\text{SIR} > 16$ [dB] ; de plus l'algorithme impliqué est dit consistant si l'écart type du SIR est inférieur à (± 10) . Une analyse de Monte Carlo a été effectuée en générant d'une façon aléatoire et uniformément des matrices \mathbf{A}_i , ce qui permet de simuler des observations $\mathbf{x}_i = \mathbf{A}_i \mathbf{s}$. Le vecteur source \mathbf{s} utilisé dans cette simulation est constitué de quatre sources non périodiques (figure 3.3 (a)) choisies dans le fichier "10 signals nonperiod" [Cichocki 2007]. Ces quatre sources remplissent bien les hypothèses de séparation exigées par toutes les méthodes que nous allons comparer.

| Méthode | SOBI | JADE ₃ | JADE | JADE ₅ | JADE ₆ | FPICA | Fast-ICA |
|---------------------------------|-------|-------------------|------|-------------------|-------------------|-------|----------|
| Mean SIRs for \mathbf{A} [dB] | 27.54 | 12.76 | 3.36 | 33.97 | 6.74 | 14.60 | 12.14 |

TABLE 3.3 – Performances (ACI, troisième simulation)

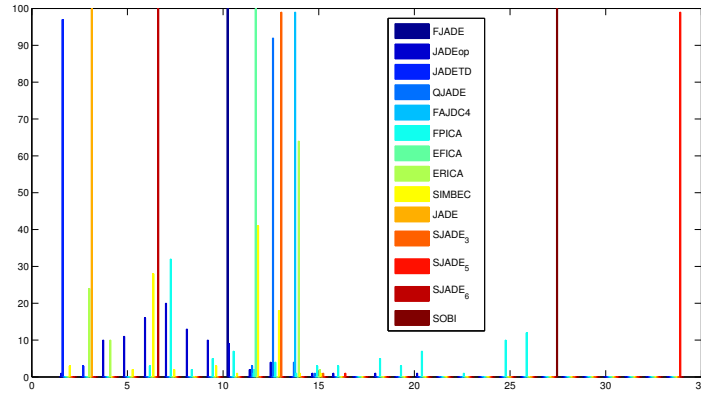


FIGURE 3.2 – La moyenne du rapport signal à interférence (SIR) pour la matrice estimée de \mathbf{A} [dB] (ACI, troisième simulation)

Les SIR sont présentés dans le tableau 3.3 et sur la figure 3.2. On constate ici que l'algorithme JADE₅ obtient de loin la meilleure séparation (figure 3.3 (a)) ; les algorithmes JADE d'ordre 3, 4 (figure 3.3 (c)) et 6 faisant moins bien que SOBI.

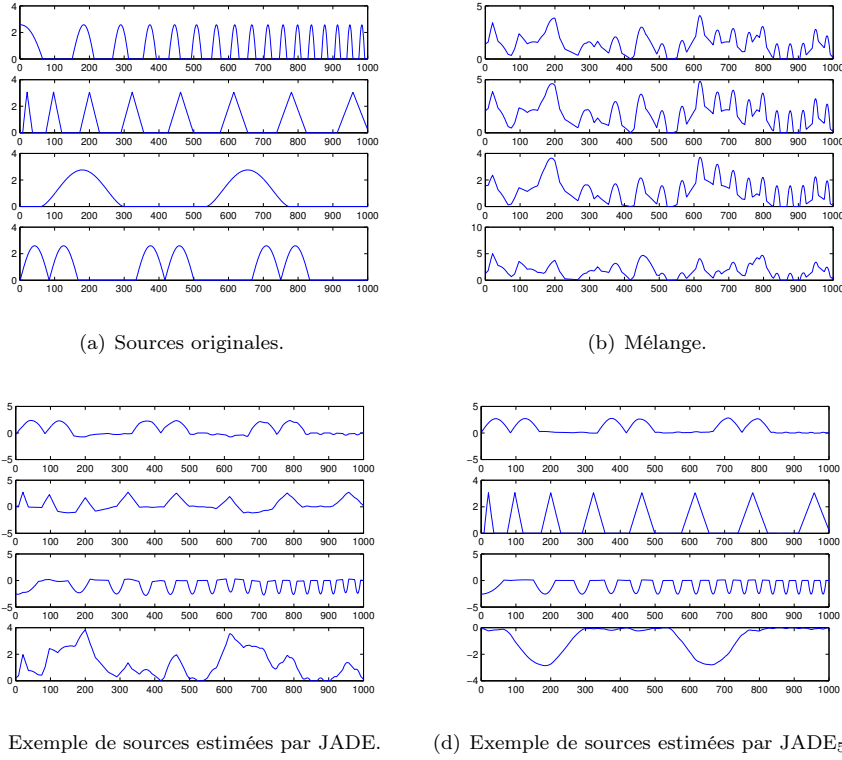


FIGURE 3.3 – ACI, troisième simulation

3.9.3 Comparaisons dans le cadre multidimensionnel (ACIM)

3.9.3.1 Première simulation

Dans une première simulation, nous considérons trois sources synthétiques, dont deux sont vectorielles de dimension 2.

$$\begin{aligned}
 \mathbf{s}_1 &= \begin{pmatrix} \exp(2(0.007t + 0.5 - \text{floor}(0.007t + 0.5)) - 1) \\ 2(0.007t + 0.5 - \text{floor}(0.007t + 0.5)) - 1 \end{pmatrix} \\
 \mathbf{s}_2 &= \begin{pmatrix} \exp(\cos(0.3t)) \\ \cos(0.3t) \end{pmatrix} \\
 \mathbf{s}_3 &= \text{square}(0.2t)
 \end{aligned} \tag{3.95}$$

On utilise un échantillonnage de 3000 points. Une analyse de Monte Carlo a été effectuée en générant d'une façon aléatoire et uniforme des matrices \mathbf{A}_i , de façon à engendrer des observations $\mathbf{x}_i = \mathbf{A}_i \mathbf{s}$. Pour obtenir une mesure de référence de la performance de chaque algorithme, nous avons utilisé des matrices de mélange carrées, sans ajouter de bruit. L'algorithme de BDC utilisé dans toutes les simulations suivantes est celui de F. Theis [Theis 2007a], avec un choix du seuil (thresholds) $\theta = 0,06$. Les performances des méthodes d'ACIM sont présentées sur la figure 3.4. Cette simulation nous montre que la méthode SJADE₆ enregistre la meilleure performance.

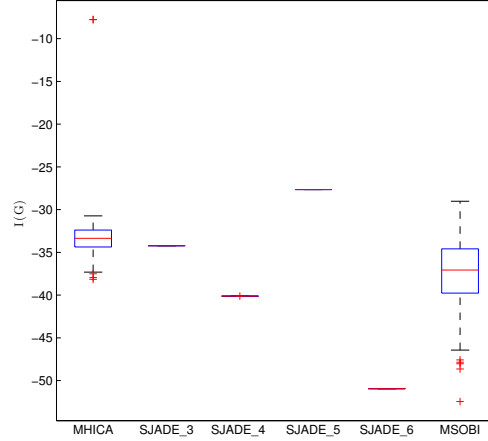


FIGURE 3.4 – Box-plot (ou boîte à moustache) de $I(\mathbf{G})$ pour les algorithmes SJADE_3 , SJADE [Theis 2007a], SJADE_5 , SJADE_6 , MHICA , [Theis 2005a] et MSOBI [Theis 2005a], pour 100 réalisations.

3.9.3.2 Seconde simulation

Voici une simulation visuelle, selon une idée introduite précédemment dans [Póczos 2005b, Theis 2007a], où le vecteur aléatoire $\mathbf{s} = (\mathbf{s}'_1, \mathbf{s}'_2, \mathbf{s}'_3, \mathbf{s}'_4)'$, de dimension 8, est formé de quatre sources \mathbf{s}_1 , \mathbf{s}_2 , \mathbf{s}_3 et \mathbf{s}_4 , de dimensions respectives 1, 3, 2 et 2. La source \mathbf{s}_1 est calculée par la formule

$$\mathbf{s}_1(\omega) = \exp(2(0.007\omega + 0.5 - \text{floor}(0.007\omega + 0.5)) - 1), \quad (3.96)$$

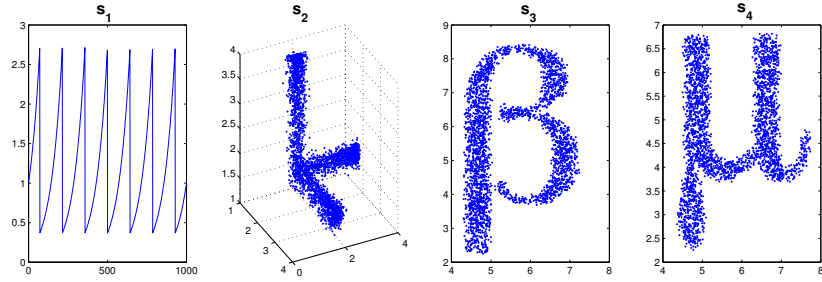
tandis que les sources \mathbf{s}_2 , \mathbf{s}_3 et \mathbf{s}_4 sont de densité uniforme sur leur support, respectivement de la forme d'un trièdre, d'un β et d'un μ . La figure 3.5 (a) représente 3000 échantillons des sources ; celles-ci sont ensuite mélangées (figure 3.5 (b)) par une matrice $\mathbf{A} \in \mathbb{R}^{8 \times 8}$, dont les coefficients sont choisis aléatoirement selon une loi uniforme sur l'intervalle $[0, 1]$ ($\mathbf{A} = \text{rand}(8)$).

Nous avons utilisé les algorithmes d'ACIM pour extraire ces quatre sources. L'indice de performance (calculé en supposant connues les dimensions des sources) est donné dans la table 3.4 pour chacune des méthodes utilisées. SJADE_5 réalise la meilleure performance.

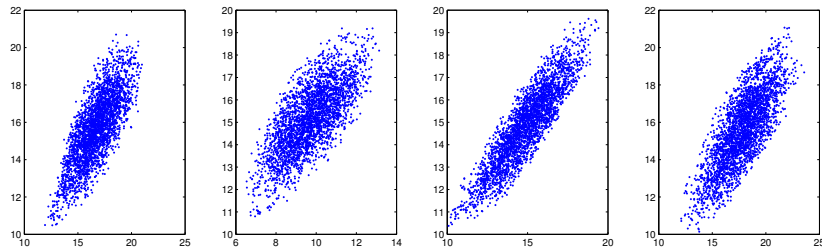
| Méthode | MSOBI | SJADE_3 | SJADE_4 | SJADE_5 | SJADE_6 | MHICA |
|----------------------------|-------|------------------|------------------|------------------|------------------|--------|
| $I(\mathbf{G})[\text{dB}]$ | -4.05 | - 26.76 | -30.10 | -32.09 | -24.21 | -24.95 |

TABLE 3.4 – Performances (ACIM, seconde simulation)

Les sources estimées sont visuellement présentées pour les méthodes SJADE_3 à SJADE_6 sur la figure 3.6. Pour tester les matrices de séparation produites par les différentes méthodes, nous traçons les profils des matrices produit de la matrice de séparation par la matrice de mélange. La figure 3.7 donne les résultats pour les méthodes MSOBI , SJADE_4 et SJADE_5 ; \mathbf{B}_2 , \mathbf{B}_4 et \mathbf{B}_5 désignant les matrices de séparations correspondantes. Clairement, les matrices \mathbf{B}_4 et \mathbf{B}_5 sont égales à permutation et matrice inversible près,



(a) Les quatre sources.



(b) Projection du mélange sur 4 plans.

FIGURE 3.5 – *ACIM, seconde simulation.*

ce qui correspond aux indéterminations de l'ACIM. Par contre, la méthode MSOBI a échoué complètement lors de cette simulation en identifiant uniquement deux composantes indépendantes de dimensions 1 et 7.

3.10 Conclusion

Nous avons montré que sous les hypothèses usuelles du problème d'ACIM, les matrices de cumulants contractées de tout ordre sont toutes bloc-diagonalisables dans la même base. Ceci permet d'envisager de résoudre le problème d'ACIM en bloc-diagonalisant conjointement un ensemble de matrices en combinant des cumulants d'ordres variés. Nous obtenons ainsi toute une famille d'algorithmes baptisés $\text{SJADE}_{r_1, \dots, r_q}$ pour Subspace JADE d'ordres r_1, \dots, r_q . Les simulations numériques montrent que le choix d'un ordre peut être crucial afin de garantir la meilleure séparation possible, ce qui prouve l'avantage qu'il y a à combiner plusieurs ordres. Par ailleurs, nous avons montré aussi que le cas le plus simple à résoudre est celui où toutes les sources vérifient la condition d'irréductibilité (3.68) et ont des dimensions deux à deux différentes ou égale à 1.

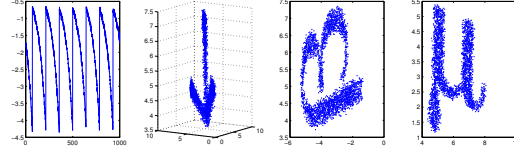
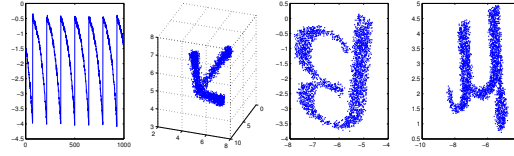
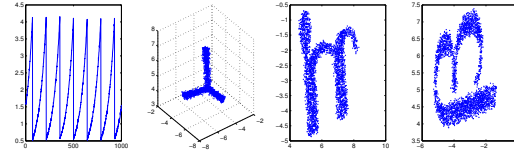
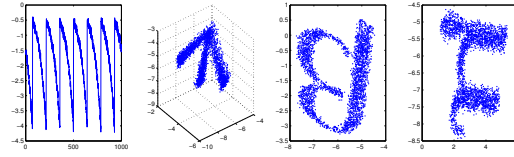
(a) Sources estimées par SJADE₃.(b) Sources estimées par SJADE₄.(c) Sources estimées par SJADE₅.(d) Sources estimées par SJADE₆.

FIGURE 3.6 – Les sources estimées (ACIM, seconde simulation)

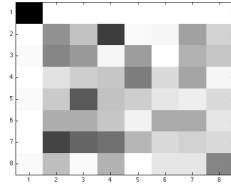
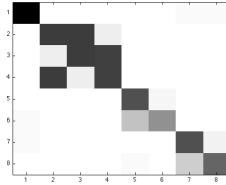
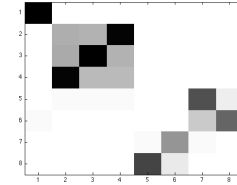
(a) $\mathbf{G}(\text{MSOBI}) = \mathbf{B}_2\mathbf{A}$;(b) $\mathbf{G}(\text{SJADE}_4) = \mathbf{B}_4\mathbf{A}$;(c) $\mathbf{G}(\text{SJADE}_5) = \mathbf{B}_5\mathbf{A}$;

FIGURE 3.7 – Test des matrices de séparation (ACIM, seconde simulation)

ACIM Robuste, par blanchiment d'ordre supérieur à 2

4.1 Introduction

En analyse en composantes indépendantes scalaires (ACI), de nombreuses méthodes de séparation comme STOTD [De Lathauwer 2001], infomax [Bell 1995], Fast-ICA [Hyvärinen 1997], SOBI [A. Belouchrani 1997], COM1 [Comon 2001], JADE [Cardoso 1993], ou encore la famille JADE_r, sont très performantes dans le cadre de mélanges linéaires non bruités. Cependant, ces méthodes peuvent être fortement perturbées par la présence d'un bruit gaussien de cohérence spatiale inconnue, car toutes nécessitent une étape de blanchiment d'ordre deux (voir [Albera 2004]). Récemment, afin de surmonter cette difficulté, plusieurs algorithmes d'ACI ont été proposés ; sans prétendre à l'exhaustivité, citons

- l'algorithme RSOBI [Belouchrani 2000], qui améliore l'algorithme SOBI [A. Belouchrani 1997], en particulier au niveau de l'étape de blanchiment ;
- une version modifiée de l'algorithme JADE, baptisée JADE cyclique [Ferreol 1997], a été proposée pour exploiter l'éventuelle cyclostationnarité des sources ;
- la méthode ICAR [Albera 2005], qui permet d'identifier la matrice de mélange sans blanchiment d'ordre deux ;
- une autre version améliorée et plus robuste de JADE baptisée Quasi-JADE [Bonhomme 2008] ;
- une extension de la méthode SOBI à l'ordre quatre, portant le nom FOBIMUM [Ferreol 2005], qui exploite les caractéristiques temporelles des sources afin de ne pas utiliser des statistiques d'ordre trop élevé ;
- la famille de méthodes BIOME [Albera 2004], qui exploite les redondances matricielles dans les matrices de cumulants (quadricovariance, hexacovariance, ...) des observations.

Signalons aussi, bien que ce problème ne soit pas abordé dans cette thèse, que FOBIMUM et BIOME permettent aussi l'identification des mélanges sous-déterminés (la dimension des sources excède éventuellement celle des observations).

Dans la même logique, en analyse en composantes indépendantes multidimensionnelles (ACIM) toutes les méthodes à blanchiment d'ordre deux seront fortement perturbées par la présence de bruit gaussien de cohérence spatiale inconnue. C'est en particulier le cas de la famille SJADE_r proposée dans le chapitre précédent. Cette raison a amené plusieurs chercheurs à étudier un nouveau problème appelé analyse en composantes non-gaussiennes (en anglais : non-gaussian component analysis) [Blanchard 2006, Theis 2006, Kawanabe 2007b, Kawanabe 2007a]. Il s'agit d'exclure la composante gaussienne indépendante (voir la définition 2) afin d'éviter toutes sortes de perturbations additionnelles. L'analyse en composantes non-gaussiennes peut ainsi être vue comme une étape de pré-traitement des données avant l'utilisation des méthodes d'ACIM.

Dans ce chapitre, nous envisagerons de résoudre le problème d'ACIM perturbé par un bruit gaussien en surpassant l'étape de blanchiment d'ordre deux, c'est-à-dire en exploitant uniquement les statistiques d'ordres supérieurs à deux.

Commençons par cette définition.

Définition 11. Soit \mathbf{x} , un vecteur aléatoire de dimension p . Nous appelons matrice de kurtosis de \mathbf{x} la matrice définie par

$$\mathcal{C}_{(4)}^{\mathbf{x}} = \mathcal{Q}_{(4)}^{\mathbf{x}}[\mathbf{1}_{p \times p}] = \mathbf{1}_{p \times p} * c_4(\mathbf{x}) = \sum_{i_1, i_2=1}^n [c_4(\mathbf{x})]_{i_1, i_2} \quad (4.1)$$

Cette matrice de taille $p \times p$ a été introduite initialement dans [Kollo 2008]. Elle n'est rien d'autre que la matrice de cumulants contractée par la matrice $\mathbf{1}_{p \times p}$ (voir la définition 3.4), autrement dit la somme des p^2 blocs matriciels de taille $p \times p$ qui constituent la matrice de cumulants d'ordre quatre $c_4(\mathbf{x})$. Elle contient donc le mélange de $\frac{p(p+1)(p+2)(p+3)}{24}$ différents cumulants scalaires $\text{cum}(x_i, x_j, x_k, x_l)$ (voir [Loperfido 2014]). Par contraste, la matrice $\mathcal{Q}_{(4)}^{\mathbf{x}}[\mathbf{I}_p]$, largement utilisée dans la séparation de sources (par exemple par la méthode FOBI [Cardoso 1989b]) ne contient le mélange que de $\frac{p^2(p+1)}{2}$ cumulants scalaires. Notons bien qu'il existe dans la littérature d'autres matrices portant le même nom (matrice de kurtosis) mais qui sont différentes (voir par exemple [Móri 1993, Tkachenko 2001, Peña 2010]).

Rappelons maintenant le modèle d'ACIM bruité.

4.2 Modèle et hypothèses pour l'ACIM en présence de bruit

$$\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{n} = \left(\sum_{i=1}^p \mathbf{A}_i \mathbf{s}_i \right) + \mathbf{n} \quad (4.2)$$

$$\text{où : } \begin{cases} n &= \sum_{j=1}^p n_j & \text{est la dimension (inconnue) du vecteur des sources,} \\ p &\in \mathbb{N} & \text{désigne le nombre (inconnu) de sources multidimensionnelles,} \\ \mathbf{A} &\in \mathbb{R}^{m \times n} & \text{représente la matrice de mélange (inconnue),} \\ \mathbf{s} &\in \mathbb{R}^n & \text{est le vecteur (inconnu) regroupant toutes les sources,} \\ \mathbf{s}_i &\in \mathbb{R}^{n_i} & \text{désigne la } i^{\text{ème}} \text{ source (inconnue),} \\ \mathbf{A}_i &\in \mathbb{R}^{m \times n_i} & \text{désigne la } i^{\text{ème}} \text{ matrice colonne de } \mathbf{A}, \\ \mathbf{x}_i &= \mathbf{A}_i \mathbf{s}_i \in \mathbb{R}^m & \text{désigne la } i^{\text{ème}} \text{ composante indépendante (inconnue),} \\ \mathbf{x} &\in \mathbb{R}^m & \text{représente le vecteur des observations (connu).} \\ \mathbf{n} &\in \mathbb{R}^m & \text{représente le vecteur bruit supposé gaussien (inconnu).} \end{cases} \quad (4.3)$$

De plus, nous adopterons les hypothèses suivantes :

- (H₁) la matrice \mathbf{A} est de rang n ,
- (H₂) les sources indépendantes sont irréductibles et non gaussiennes,
- (H₃) \mathbf{s} et \mathbf{n} sont indépendantes,
- (H₄) la matrice de kurtosis $\mathcal{C}_{(4)}^{\mathbf{x}}$ est semi-définie positive ou semi-définie négative et de rang n ,
- (H₅) une seule source (scalaire) peut avoir un cumulant d'ordre $r(r > 4)$ nul.

4.3 Une méthode de séparation robuste au bruit additif gaussien

4.3.1 Un blanchiment d'ordre quatre

Sous l'hypothèse d'indépendance (H₃), nous avons

$$c_4(\mathbf{x}) = c_4(\mathbf{A}\mathbf{s}) = (\mathbf{A} \otimes \mathbf{A})c_4(\mathbf{s})(\mathbf{A} \otimes \mathbf{A})' + c_4(\mathbf{n}) = (\mathbf{A} \otimes \mathbf{A})c_4(\mathbf{s})(\mathbf{A} \otimes \mathbf{A})', \quad (4.4)$$

puisque $c_4(\mathbf{n}) = \mathbf{0}$ (\mathbf{n} est gaussien).

Sous l'hypothèse (H_2) , la matrice de kurtosis de \mathbf{x} s'exprime directement par la formule (3.51) page 50 :

$$\mathcal{C}_{(4)}^{\mathbf{x}} = \mathcal{Q}_{(4)}^{\mathbf{x}}[\mathbf{1}_{m \times m}] = \mathbf{A} \text{ bdiag} \left(\mathcal{Q}_{(4)}^{s_1}[\mathbf{A}'_1 \mathbf{1}_{m \times m} \mathbf{A}_1], \dots, \mathcal{Q}_{(4)}^{s_p}[\mathbf{A}'_p \mathbf{1}_{m \times m} \mathbf{A}_p] \right) \mathbf{A}', \quad (4.5)$$

où la matrice \mathbf{A} est partitionnée comme on l'a décrit à la sous-section 3.2.4 page 41. Nous pouvons donc écrire,

$$\mathcal{C}_{(4)}^{\mathbf{x}} = \mathbf{A} \Delta \mathbf{A}', \quad (4.6)$$

où Δ est la matrice bloc-diagonale de taille $n \times n$

$$\Delta = \text{bdiag} \left(\mathcal{Q}_{(4)}^{s_1}[\mathbf{A}'_1 \mathbf{1}_{m \times m} \mathbf{A}_1], \dots, \mathcal{Q}_{(4)}^{s_p}[\mathbf{A}'_p \mathbf{1}_{m \times m} \mathbf{A}_p] \right). \quad (4.7)$$

Sous l'hypothèse (H_4) , nous supposons sans perte de généralité que $\mathcal{C}_{(4)}^{\mathbf{x}}$ est semi-définie positive (dans le cas contraire on considère $-\mathcal{C}_{(4)}^{\mathbf{x}}$). En posant $\tilde{\mathbf{A}} = \mathbf{A} \Delta^{\frac{1}{2}}$, il vient

$$\mathcal{C}_{(4)}^{\mathbf{x}} = \tilde{\mathbf{A}} \tilde{\mathbf{A}}' = \mathbf{U} \underbrace{\mathbf{D} \mathbf{V} \mathbf{V}' \mathbf{D}}_{=\mathbf{I}_n} \mathbf{U}' = \mathbf{U} \mathbf{D}^2 \mathbf{U}', \quad (4.8)$$

où $\mathbf{U} \mathbf{D} \mathbf{V}$ est une décomposition en valeurs singulières (DVS) de $\tilde{\mathbf{A}}$, de sorte que \mathbf{U} est une matrice semi-orthogonale de dimension $m \times n$, \mathbf{D} une matrice diagonale de dimension $n \times n$ et \mathbf{V} une matrice orthogonale de dimension $n \times n$.

On voit ainsi que l'équation (4.8) est une décomposition en valeurs singulières de la matrice de kurtosis $\mathcal{C}_{(4)}^{\mathbf{x}}$, de sorte que les matrices \mathbf{U} et \mathbf{D} pourront effectivement être estimées avec la seule connaissance de $\mathcal{C}_{(4)}^{\mathbf{x}}$. Ainsi, de façon similaire au blanchiment d'ordre deux (précédemment décrit à la sous-section 3.2.3 page 41), nous pouvons proposer un blanchiment des observations à l'ordre quatre en posant $\mathbf{W} = \mathbf{D}^{-1} \mathbf{U}'$ et $\mathbf{z} = \mathbf{W} \mathbf{x}$. Autrement dit, la matrice de kurtosis $\mathcal{C}_{(4)}^{\mathbf{x}}$ joue dans ce cas un rôle similaire à celui de la matrice de covariance lors d'un blanchiment ordinaire d'ordre deux.

Remarque 33 (sur l'hypothèse (H_4)). Dans le cas ACI, nous pouvons montrer facilement à partir de (4.7) (ou voir par exemple [Kollo 2008]) que

$$\Delta = \text{diag} \left(\sum_{i=1}^m \ell_i \right) \begin{pmatrix} c_4(s_1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & c_4(s_n) \end{pmatrix} \text{diag} \left(\sum_{i=1}^m \ell_i \right) \quad (4.9)$$

$$= \text{diag} \left(\sum_{i=1}^m \ell_i \right)^2 \text{diag} (c_4(s_1), c_4(s_2), \dots, c_4(s_n)), \quad (4.10)$$

où ℓ_i est la i -ème ligne de la matrice \mathbf{A} et $c_4(s_i)$ est le kurtosis de la source scalaire s_i . Ce qui veut dire que l'hypothèse (H_4) est équivalente dans cette circonstance au fait que les kurtosis des sources sont non nuls et ont le même signe. Notre hypothèse (H_4) apparaît ainsi comme une extension au cas multidimensionnel de la condition de signe des kurtosis classiquement supposée dans le cas scalaire. Dans le cas ACIM, nous pouvons seulement écrire

$$\Delta = \begin{pmatrix} \mathcal{Q}_{(4)}^{s_1}[\mathbf{A}'_1 \mathbf{1}_{n \times n} \mathbf{A}_1] & \dots & \mathbf{0}_{n_1 \times n_p} \\ \vdots & \ddots & \vdots \\ \mathbf{0}_{n_p \times n_1} & \dots & \mathcal{Q}_{(4)}^{s_p}[\mathbf{A}'_p \mathbf{1}_{n \times n} \mathbf{A}_p] \end{pmatrix}. \quad (4.11)$$

4.3.2 Estimation de la matrice orthogonale manquante

Après le blanchiment d'ordre quatre, le modèle (4.2) se transforme en

$$\mathbf{z} = \mathbf{W}\mathbf{x} = \mathbf{W}\mathbf{A}\mathbf{s} + \mathbf{W}\mathbf{n} = \mathbf{V}\mathbf{s} + \mathbf{W}\mathbf{n} \quad (4.12)$$

Il nous reste alors à estimer la matrice orthogonale $\mathbf{V} = \mathbf{W}\mathbf{A}$ pour calculer la matrice de mélange \mathbf{A} (aux indéterminations habituelles près).

Pour cela, nous allons procéder par bloc-diagonalisation de matrices de cumulants contractées, comme on l'a présenté au chapitre précédent. La formule (4.12) montre que les matrices de cumulants d'ordre $r > 2$ de \mathbf{z} s'écrivent

$$c_r(\mathbf{z}) = c_r(\mathbf{W}\mathbf{x}) \underset{(H_3)}{=} c_r(\mathbf{V}\mathbf{s}) + c_r(\mathbf{W}\mathbf{n}) = c_r(\mathbf{V}\mathbf{s}). \quad (4.13)$$

Nous pouvons donc encore utiliser notre théorème de bloc-diagonalisation (théorème 22 page 48), en particulier pour les ordres 5 et 6, les seuls utilisés dans les applications présentées plus loin. Pour tout $\mathbf{M}^{(3)} \in \mathbb{R}^{n^2 \times n}$ et tout $\mathbf{M}^{(4)} \in \mathbb{R}^{n^2 \times n^2}$, il vient

$$\mathcal{Q}_{(5)}^{\mathbf{z}}[\mathbf{M}^{(3)}] = \mathbf{V} \text{bdiag} \left(\mathcal{Q}_{(5)}^{\mathbf{s}_1} \left[(\mathbf{V}'_1 \otimes \mathbf{V}'_1) \mathbf{M}^{(3)} \mathbf{V}_1 \right], \dots, \mathcal{Q}_{(5)}^{\mathbf{s}_p} \left[(\mathbf{V}'_p \otimes \mathbf{V}'_p) \mathbf{M}^{(3)} \mathbf{V}_p \right] \right) \mathbf{V}' \quad (4.14)$$

$$\begin{aligned} \mathcal{Q}_{(6)}^{\mathbf{z}}[\mathbf{M}^{(4)}] &= \mathbf{V} \text{bdiag} \left(\mathcal{Q}_{(6)}^{\mathbf{s}_1} \left[(\mathbf{V}'_1 \otimes \mathbf{V}'_1) \mathbf{M}^{(4)} (\mathbf{V}_1 \otimes \mathbf{V}_1) \right], \dots, \right. \\ &\quad \left. \dots, \mathcal{Q}_{(6)}^{\mathbf{s}_p} \left[(\mathbf{V}'_p \otimes \mathbf{V}'_p) \mathbf{M}^{(4)} (\mathbf{V}_p \otimes \mathbf{V}_p) \right] \right) \mathbf{V}' \end{aligned} \quad (4.15)$$

où $\mathbf{V} = \begin{pmatrix} \mathbf{V}_1 & \dots & \mathbf{V}_p \end{pmatrix}$ désigne le partitionnement idoine.

Il est ainsi clair qu'on peut espérer calculer \mathbf{V} à l'aide d'une bloc-diagonalisation conjointe (BDC) de matrices de cumulants contractées $\mathcal{Q}_{(r)}^{\mathbf{z}}[\mathbf{M}]$.

4.3.3 RSJADE_r : une famille d'algorithmes pour l'ACIM robuste

Voici maintenant le synoptique de l'algorithme que nous baptisons RSJADE_r (pour Robust Subspace JADE d'ordre r) qui résout le problème d'ACIM bruité sous les hypothèses précédentes en blanchissant les observations à l'ordre 4.

Algorithme 34 (RSJADE_r).

1. **blanchiment à l'ordre 4 des observations :**

- (a) $\mathbf{x} = \mathbf{x} - \mathbb{E}(\mathbf{x})$ (centrage);
- (b) **estimation de la matrice de kurtosis :** calcul de la matrice de kurtosis $C_{(4)}^{\mathbf{x}}$; décomposition en valeurs singulières de $C_{(4)}^{\mathbf{x}}$; si elle n'est pas définie positive, considérer $-C_{(4)}^{\mathbf{x}}$ (hypothèse (H_4)); estimation du nombre n de sources par le nombre de valeurs propres non nulles de $C_{(4)}^{\mathbf{x}}$; obtention de la matrice de blanchiment $\mathbf{W} : n \times m$;
- (c) $\mathbf{z} = \mathbf{W}\mathbf{x}$ (blanchiment d'ordre 4);

2. **construction d'un ensemble de matrices de cumulants contractées :** par exemple l'ensemble

$$\mathcal{M}^{(r)} = \left\{ \mathcal{Q}_{(r)}^{\mathbf{z}} \left[\mathbf{E}_{i_1, i_2, \dots, i_{r-2}} \right] \quad ; \quad 1 \leq i_1, i_2, \dots, i_{r-2} \leq n \right\} \quad (4.16)$$

de, a priori, n^{r-2} matrices, calculé à partir d'une simple estimation empirique de la matrice de cumulants $c_r(\mathbf{z})$. Cette étape utilise les formules développées dans le chapitre 2.

3. **bloc-diagonalisation conjointe (BDC) :** estimation de la matrice orthogonale \mathbf{V}' qui bloc-diagonalise conjointement $\mathcal{M}^{(r)}$.

4. estimation des sources :

(a) $\mathbf{B} = \mathbf{V}'\mathbf{W}$ (matrice de séparation) ;

(b) $\mathbf{s} = \mathbf{B}\mathbf{x}$.

4.4 Petite simulation et discussion

Nous allons tester la méthode précédente sur ce mélange bruité synthétique,

$$\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{n} = \mathbf{A}_1\mathbf{s}_1 + \mathbf{A}_2\mathbf{s}_2 + \mathbf{n}, \quad (4.17)$$

où :

$$\mathbf{A} = \begin{pmatrix} 0.8255 & 0.4373 & 0.9321 \\ 0.4880 & 0.0884 & 0.0006 \\ 0.7110 & 0.4787 & 0.1534 \end{pmatrix} \quad (\text{matrice quelconque}) \quad (4.18)$$

$$\mathbf{s} = \begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{pmatrix} \quad \text{avec : } \begin{cases} \mathbf{s}_1 = \exp(2(0.007 t + 0.5 - \text{floor}(0.007 t + 0.5)) - 1) \\ \mathbf{s}_2 = \begin{pmatrix} \sin(0.3 t) \\ \text{square}(0.1 t) \end{pmatrix} \end{cases} \quad (4.19)$$

et \mathbf{n} un vecteur gaussien de matrice de covariance de la forme

$$\text{cov}(\mathbf{n}) = \mathbf{R}_n = \sigma^2 \begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{pmatrix} \quad (4.20)$$

où σ^2 est la puissance du bruit et ρ son coefficient de corrélation.

Nous pouvons alors calculer¹ les matrices de cumulants d'ordre 4 des sources \mathbf{s}_1 et \mathbf{s}_2 ; on trouve respectivement :

$$c_4(\mathbf{s}_1) = -0.1333 \quad \text{et} \quad c_4(\mathbf{s}_2) = \begin{pmatrix} -0.3751 & -0.1773 & -0.1773 & -0.0905 \\ -0.1773 & -0.0905 & -0.0905 & -0.4253 \\ -0.1773 & -0.0905 & -0.0905 & -0.4253 \\ -0.0905 & -0.4253 & -0.4253 & -1.9997 \end{pmatrix}. \quad (4.21)$$

Nous calculons aussi les matrices de covariance et de kurtosis (définition 11) du vecteur source \mathbf{s} :

$$\text{cov}(\mathbf{s}) = \mathbf{R}_s = \begin{pmatrix} 0.3334 & 0.0016 & 0.0034 \\ 0.0016 & 0.5002 & 0.2127 \\ 0.0034 & 0.2127 & 1.0002 \end{pmatrix} ; \quad \mathcal{C}_{(4)}^s = \begin{pmatrix} -0.1339 & -0.0053 & -0.0114 \\ -0.0053 & -0.8257 & -0.7897 \\ -0.0114 & -0.7897 & -2.9595 \end{pmatrix}. \quad (4.22)$$

Enfin, supposant connu uniquement le vecteur des observations \mathbf{x} , nous utilisons différentes méthodes de séparation pour tenter de retrouver les sources \mathbf{s}_1 et \mathbf{s}_2 (aux indéterminations près, comme toujours).

La figure (4.1) montre les performances des méthodes SJADE [Theis 2007a], SJADE₆ et RSJADE₆, pour $\rho = 0.9$ en fonction du rapport signal sur bruit RSB qui varie de 0dB à 20dB. Nous pouvons constater la supériorité de RSJADE₆ sur SJADE₄ et SJADE₆, qui sont deux méthodes utilisant un blanchiment d'ordre deux.

La famille RSJADE_r est une extension multidimensionnelle de la méthode d'ACI que nous avons introduite dans [Ould-Baba 2011]. Celle-ci est très robuste et a l'avantage de résoudre le problème d'ACI

1. Tous les calculs présentés dans cette section sont effectués sur un échantillon de 3000 réalisations.

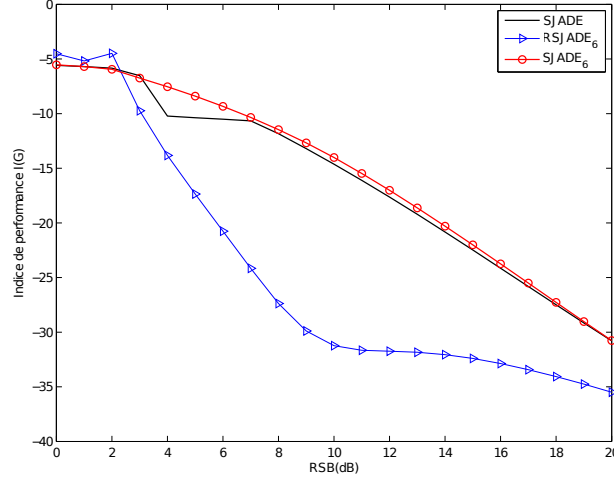


FIGURE 4.1 – ACIM, simulation, indice de performance $I(G)$, en fonction du RSB.

sous une contrainte sur la matrice de mélange \mathbf{A} vérifiable en pratique, contrairement à d'autres méthodes de cette catégorie telles que ICAR [Albera 2005], FOBUM [Ferreol 2005] et BIOME [Albera 2004], qui supposent toutes que la matrice de mélange ne contient aucun élément nul, chose qu'on peut pas tester apriori. En effet, notre méthode fonctionne sous l'hypothèse (H_4) qui demande seulement que $\text{rang}(C_{(4)}^{\mathbf{x}}) = \text{rang}(\mathbf{A})$. Cependant, une telle hypothèse exclut de fait le traitement des mélanges sous-déterminés.

4.5 Si la matrice de kurtosis n'est ni positive ni négative

Dans l'ACI, on sait que le cas où les kurtosis des sources n'ont pas le même signe est une situation difficile où aucune des méthodes d'ordres supérieurs (ICAR, FOBUM et RSJADE_r) ne fonctionnent. Notre hypothèse (H_4) est l'analogue multidimensionnel de cette condition de même signe des kurtosis. Il pourrait donc sembler illusoire de s'en passer.

Cependant, si (H_4) est en défaut, nous pouvons contourner cette difficulté en changeant de matrice de blanchiment d'ordre quatre. Il suffit pour cela de choisir une matrice de cumulants contractée $\mathcal{Q}_{(4)}^{\mathbf{x}}[\mathbf{M}]$ qui soit semi-définie positive. Cela revient à trouver \mathbf{M} , matrice contractante, telle que

$$\mathcal{Q}_{(4)}^{\mathbf{x}}[\mathbf{M}] = \sum_{i_1, i_2=1}^m m_{i_1, i_2} [c_4(\mathbf{x})]_{i_1, i_2} \geq 0, \quad (4.23)$$

problème résolu facilement et efficacement via des algorithmes tels que [Tong 1992, Zaidi 2005].

Vérifions-le sur un exemple. Soit un mélange de 4 sources indépendantes, deux sub-gaussiennes de kurtosis respectifs $k_{s_1} = -1.9969$ et $k_{s_2} = -1.9993$, deux super-gaussiennes de kurtosis respectifs $k_{s_3} = 4.6591$ et $k_{s_4} = 5.5660$. Les simulations numériques (voir la figure 4.2) ont montré que dans une telle situation, une variante de l'algorithme RSJADE_r à blanchiment d'ordre quatre obtenu à partir de la matrice $\mathcal{Q}_{(4)}^{\mathbf{x}}[\mathbf{M}]$, où \mathbf{M} a été calculée via l'algorithme [Tong 1992], permet de résoudre le problème. La

matrice de mélange choisie pour cette simulation est

$$\mathbf{A} = \begin{pmatrix} 0.5037 & 0.1933 & 0.6561 & 0.8455 \\ 0.7517 & 0.5153 & 0.2603 & 0.9610 \\ 0.9433 & 0.8395 & 0.6930 & 0.4016 \\ 0.6910 & 0.5627 & 0.9629 & 0.3907 \end{pmatrix}. \quad (4.24)$$

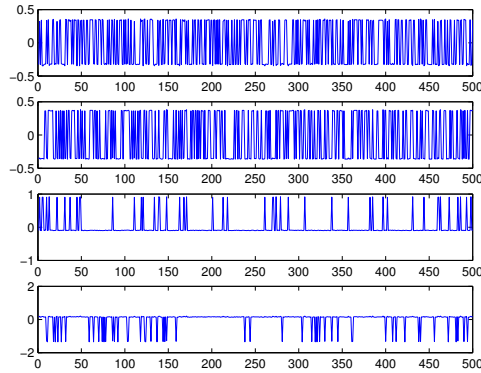
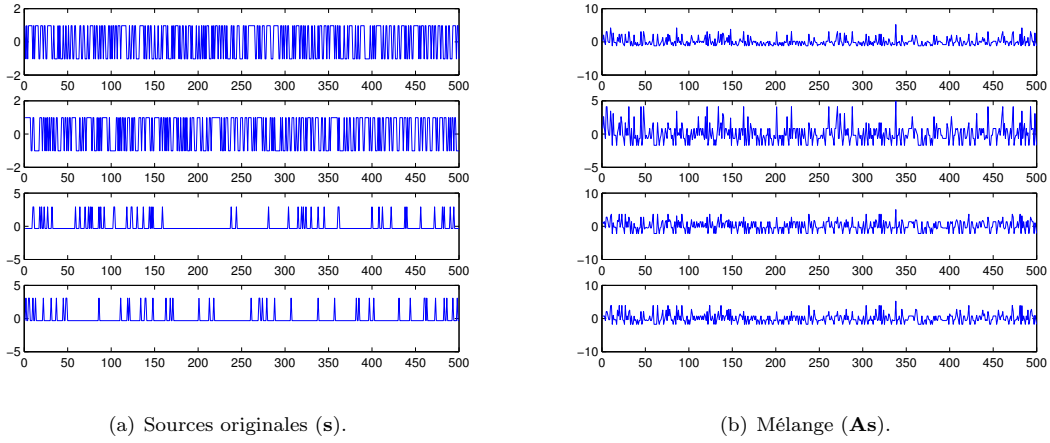


FIGURE 4.2 – *ACI, simulation (kurtosis de signes différents).*

Une autre possibilité consiste à utiliser un blanchiment d'ordre 3, dans le cas où la matrice de skewness

$$\mathcal{C}_{(3)}^{\mathbf{x}} = \mathcal{Q}_{(3)}^{\mathbf{x}}[\mathbf{1}_p] = \mathbf{1}_p * c_3(\mathbf{x}) = \sum_i^m [c_3(\mathbf{x})]_i = \mathbb{E}((\mathbf{x}'\mathbf{1}_p)\mathbf{x}\mathbf{x}'), \quad (4.25)$$

vérifie l'hypothèse (H_4). Nous pouvons ainsi engendrer d'une manière analogue à RSJADE_r, une nouvelle famille de méthodes à blanchiment d'ordre 3 et ainsi de suite.

Remarque 35 (pourquoi pas une extension multidimensionnelle de la méthode ICAR [Albera 2005]?).

Il est clair qu'à partir de l'équation

$$c_4(\mathbf{x}) = \begin{pmatrix} \mathbf{A}_1 \otimes \mathbf{A}_1 & \cdots & \mathbf{A}_p \otimes \mathbf{A}_p \end{pmatrix} \text{bdiag}(c_4(\mathbf{s}_1), \dots, c_4(\mathbf{s}_p)) \begin{pmatrix} \mathbf{A}_1 \otimes \mathbf{A}_1 & \cdots & \mathbf{A}_p \otimes \mathbf{A}_p \end{pmatrix}', \quad (4.26)$$

nous pourrions envisager de résoudre le problème d'ACIM (bruité), par la méthode d'ACI, ICAR [Albera 2005] puis chercher une permutation qui rangera les sorties (sources scalaires) en fonction de leurs dépendance. Malheureusement cela est très difficile à cause des points suivants :

- une source irréductible \mathbf{s}_i peut vérifier la condition (3.68), c-à-d. que $\text{rang}(c_4(\mathbf{s}_i)) = \frac{n_i(n_i+1)}{2}$, ce qui va impliquer que $\text{rang}(c_4(\mathbf{x})) = r > n$; ICAR va donc chercher r sources scalaires au lieu de n ;
- ICAR ne marche que sous l'incontournable hypothèse $c_4(\mathbf{x}) > 0$ (ou $-c_4(\mathbf{x}) > 0$); ceci implique que les matrices de cumulants des sources $c_4(\mathbf{s}_i)$ doivent être définies positives simultanément qui est une hypothèse très forte pour un algorithme qui cherche à capturer des sources dépendantes.

4.6 Applications aux signaux électrocardiogrammes

Le cœur est un muscle qui joue le rôle d'une pompe maintenant la circulation du sang dans les vaisseaux de l'organisme. Un signal ÉlectroCardioGramme (ECG) est l'enregistrement de l'activité électrique intrinsèque du cœur via des électrodes placées sur la peau de l'individu, le tissu humain étant conducteur naturel. Cet outil de diagnostic est devenu au fil des décennies indispensable pour l'examen cardiologique, à tel point que certains le considèrent actuellement comme faisant partie de l'examen médical ordinaire.

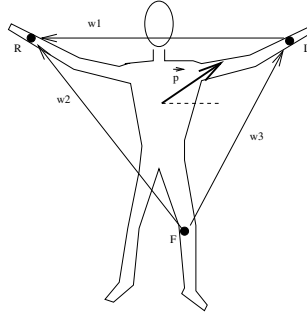


FIGURE 4.3 – *Electrodes disposées en triangle d'Einthoven.*

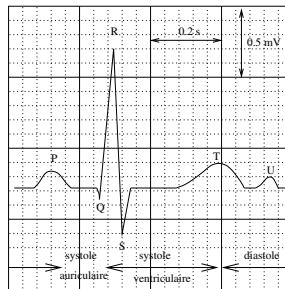


FIGURE 4.4 – *Un signal ECG typique.*

Le premier ECG a été réalisé en 1901 par Einthoven à l'aide d'un dispositif relativement lourd (600 kg de matériel, galvanomètre et bassine d'eau pour augmenter la conductivité du corps au niveau des électrodes).

Il décrit les 5 déflexions majeures, à savoir P, Q, R, S et T (que nous illustrons dans la figure (4.4)), lui valant le prix Nobel de physiologie en 1924. Sur la figure (4.3), on a représenté schématiquement la disposition des trois électrodes utilisées par Einthoven. Ces lettres sont donc caractéristiques de l'activité électrique du cœur. Ainsi,

- l'onde P correspond à la dépolarisation des cellules des oreillettes, ce qui entraîne leur contraction. L'onde P dure environ 90 ms.
- le complexe QRS correspond à la dépolarisation des cellules des ventricules, ce qui entraîne aussi leur contraction. Le QRS dure environ 90 ms, soit significativement moins que la durée de la systole ventriculaire (environ 0.3 s).
- l'onde T correspond à la repolarisation des ventricules.

4.6.1 Interprétation dipolaire

Les électrodes de l'ECG vont recueillir la somme des activités électriques de cellules cardiaques. Il semble que les praticiens résument l'activité électrique du cœur par un dipôle centré sur le cœur et de moment $\vec{p}(t)$, dont la direction et l'intensité varient de façon (presque) périodique au cours du temps. Ils disent que chaque dérivation (par exemple w_1 sur la figure 4.3) mesure la projection de $\vec{p}(t)$ dans sa direction (ici \overrightarrow{LR} pour la dérivation w_1). Ceci s'explique très simplement par :

$$w_1(t) = V(R) - V(L) \quad (4.27)$$

où $V(R)$ et $V(L)$ sont respectivement les potentiels aux points R et L . Les premières dérivations $w_1(t)$, $w_2(t)$ et $w_3(t)$ sont des dérivations bipolaires pour lesquelles trois électrodes sont utilisées, une au poignet de chaque bras (gauche L et droit R) et la troisième à la cheville gauche (F), voir (4.3). Les appareils modernes sont miniaturisés, portatifs, et utilisent 10 électrodes (6 sur le thorax et une sur chacun des membres). Le principe reste cependant le même, et pour l'illustrer, on retiendra pour l'instant le schéma simplifié de la figure (4.3).

4.6.2 Problème de l'extraction du ECG de fœtus

L'enregistrement du signal ECG n'est pas toujours facile. Par exemple, un problème classique lié à cet enregistrement est rencontré durant la grossesse, où il est très utile de pouvoir accéder à l'ECG du fœtus. Il s'agit du problème de l'extraction de battements cardiaques du fœtus à partir d'une mesure des signaux ECG enregistrés via des électrodes placées sur la peau de la mère. Il se trouve que le signal ECGF (ECG du fœtus) a une puissance très faible par rapport à la puissance de l'ECGM (ECG maternelle). De plus, il y a plusieurs autres signaux parasites qui vont venir s'ajouter aux mélanges de ces deux contributions, par exemple : le bruit intrinsèque de l'enregistreur, bruit de contact d'électrodes, la respiration de la mère, l'activité de l'estomac. La situation est encore bien pire pendant les contractions utérines de la mère. En effet, pendant ces contractions, l'enregistrement de signaux ECGF peut être perturbé par d'autres signaux électrophysiologiques tel que les signaux électromyogrammes (EMG), qui sont dus au muscle utérin plutôt qu'au cœur du fœtus. Or la réponse du cœur de fœtus pendant ces contractions est un indicateur très important de sa santé.

4.6.3 Quelques solutions du problème de l'extraction d'activité fœtale

Le problème de l'extraction des activités électriques fœtales et maternelles est formalisé pour la première fois comme un problème de séparation aveugle de sources d'un mélange linéaire instantané bruité dans [De Lathauwer 1995], où une solution a été proposée en s'appuyant sur les travaux présentés dans [De Lathauwer 1994] qui exploitent les statistiques d'ordres supérieurs. Plusieurs autres solutions ont été ensuite obtenues en s'appuyant sur d'autres approches ACI; nous citons par exemple les références

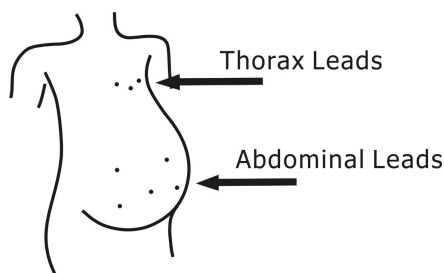


FIGURE 4.5 – Emplacement des capteurs, pour un système à 8 canaux d'acquisition des signaux ECG selon [De Lathauwer 1995], 5 sur la région abdominale et 3 sur la cage thoracique.

[Choi 2000, Zarzoso 2009, Ye 2008, Vigneron 2003, Azzerboni 2005, Boustany 2005, Marossero 2003]. Cependant, comme il n'est pas raisonnable de supposer l'activité d'une partie du coeur complètement indépendante de l'activité d'une autre partie (voir [Chawla 2009a]), il est devenu indispensable de passer à des extensions multidimensionnelles de l'ACI. Le problème de l'extraction d'activité fœtale refait surface sous ce nouvel angle d'attaque. Tout d'abord dans [Cardoso 1998] où Cardoso choisit trois mesures ECG prises sur la région abdominale accessible dans la base de données [De Moor 1997]; l'algorithme employé pour cette étude est JADE [Cardoso 1993]. Cardoso a pu extraire les deux sous-espaces ECGM et ECGF en regroupant les composantes scalaires fournies par JADE pour engendrer une solution ACIM. Depuis, diverses méthodes d'analyse en composantes indépendantes multidimensionnelles ont été testées pour améliorer la solution du problème. Nous citons en particuliers les références [Theis 2005a, Meinecke 2002, Sharma 2006, Li 2010, Yonghong 2012, Camargo-Olivares 2011]. Par ailleurs, récemment, l'ACIM a été aussi utilisée pour la correction de signaux ECG eux-même dans les études présentées dans les références [Chawla 2009a, Chawla 2009b].

4.6.4 Résolution du problème via nos algorithmes

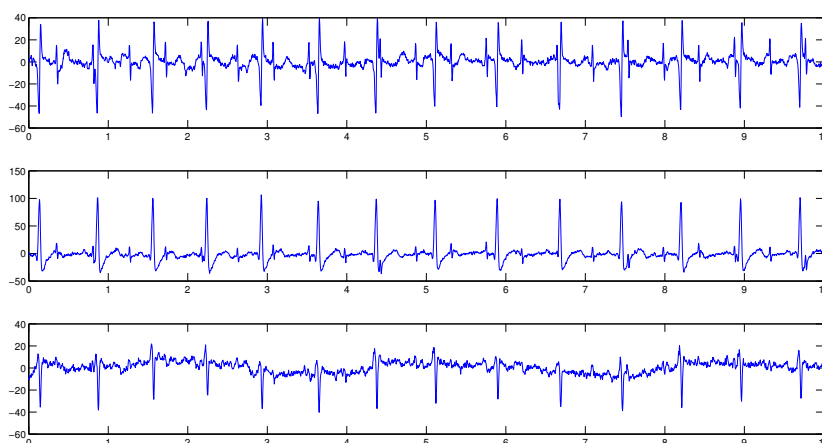


FIGURE 4.6 – les 3 mesures sélectionnés à partir de la base de données [De Moor 1997].

Nous allons maintenant utiliser les méthodes d'analyse en composantes indépendantes multidimensionnelles que nous avons développées tout-au-long de cette thèse pour résoudre ce problème. Pour cela, nous avons choisi, comme Cardoso, trois mesures ECG prises uniquement sur la région abdominale d'une mère enceinte, accessible dans la base de données [De Moor 1997]. Il s'agit de 2500 points ECG, la fréquence d'échantillonnage est de 250Hz, voir la figure (4.6). Les signaux observés sont le mélange des rythmes cardiaques de la mère et de son fœtus. Notre objectif est de séparer ce mélange en deux composantes multidimensionnelles représentant respectivement les sous espaces ECGM et ECGF.

Nous allons utiliser successivement la famille de méthodes à blanchiment d'ordre deux $\text{SJADE}_r(r = 3, \dots, 6)$ (voir l'algorithme 24 page 51) et la famille à blanchiment d'ordre quatre $\text{RSJADE}_r(r = 5, 6)$ (algorithme 34 page 70). Les deux sources multidimensionnelles ECGM (2D), ECGF (1D) estimées respectivement par ces deux familles de méthodes sont présentées dans les figures (4.7) et (4.8).

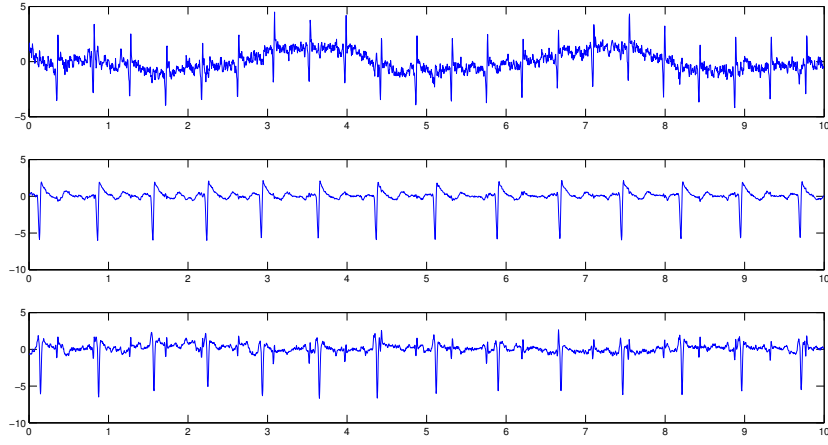
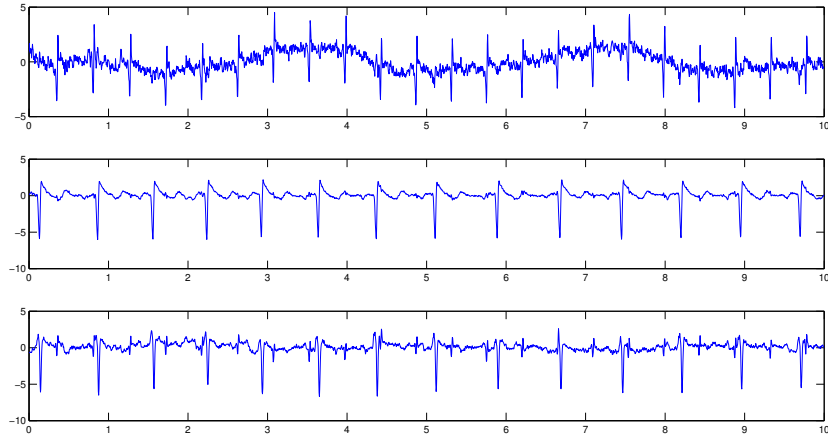
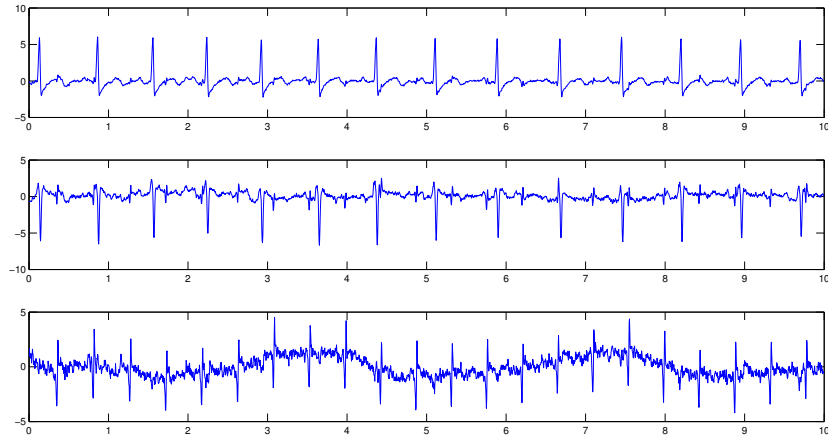
(a) SJADE₄.(b) SJADE₅(c) SJADE₆.

FIGURE 4.7 – Les contributions ECGM (de dimension 2) et ECGF (de dimension 1) estimées par $SJADE_r$, ($r \in \{4, 5, 6\}$) à blanchiment d'ordre 2.

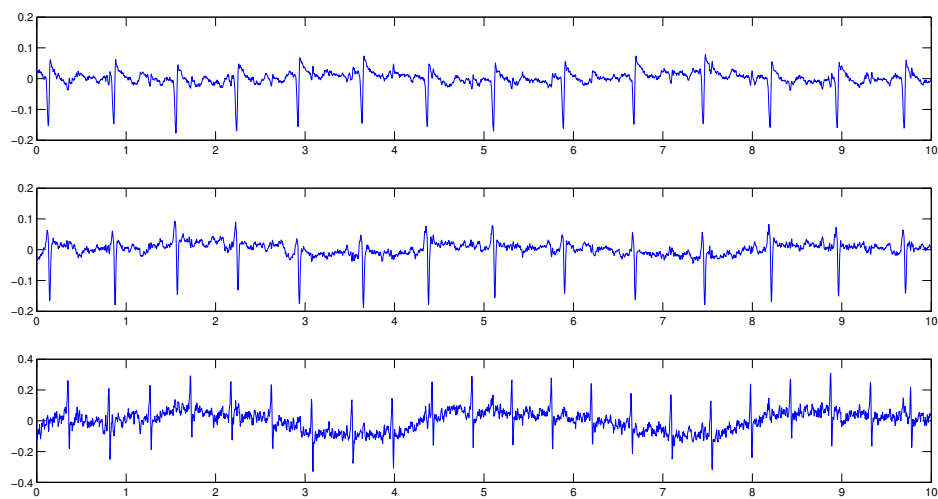
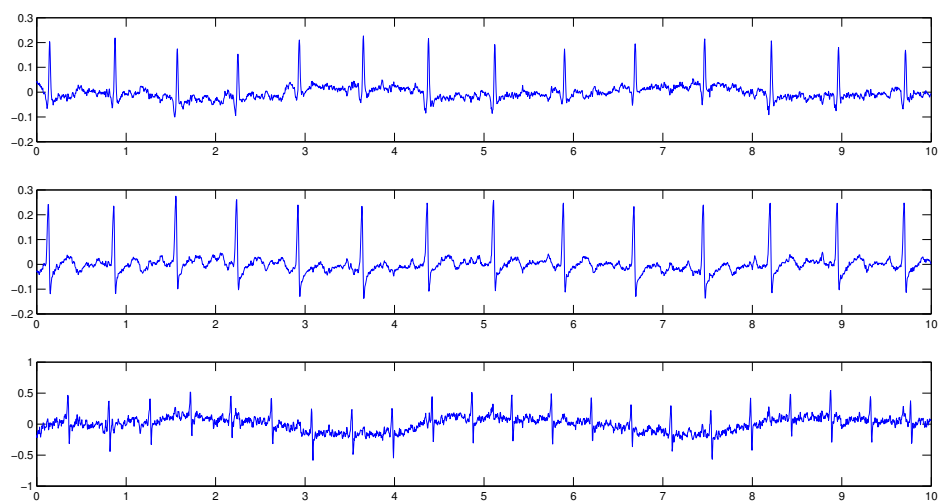
(a) RSJADE₅.(b) RSJADE₆

FIGURE 4.8 – Les contributions ECGM (de dimension 2) et ECGF (de dimension 1) estimées par la famille $RSJADE_r$, ($r \in \{5, 6\}$) à blanchiment d'ordre 4.

Ces deux contributions ont été correctement séparées. Ceci est confirmé par le nombre de pic R (voir (4.4)) qui distingue le cycle cardiaque du fœtus à celui de la mère. Dans la suite, nous donnons les matrices des séparations estimées, nous notons par \mathbf{B}_r celle de la famille de méthodes SJADE_r et par \mathbf{B}_r^4 celle de la famille de méthode RSJADE_r à blanchiment d'ordre 4.

$$\begin{aligned}\mathbf{B}_3 &= \begin{pmatrix} 0.1018 & 0.0428 & 0.1606 \\ 0.0495 & -0.0394 & -0.0177 \\ -0.0699 & -0.0352 & 0.0601 \end{pmatrix}; & \mathbf{B}_4 &= \begin{pmatrix} 0.1027 & 0.0442 & 0.1594 \\ 0.0515 & -0.0385 & -0.0159 \\ 0.0671 & 0.0345 & -0.0638 \end{pmatrix} \\ \mathbf{B}_5 &= \begin{pmatrix} 0.1028 & 0.0451 & 0.1588 \\ 0.0534 & -0.0376 & -0.0144 \\ -0.0653 & -0.0342 & 0.0656 \end{pmatrix}; & \mathbf{B}_6 &= \begin{pmatrix} 0.0570 & -0.0360 & -0.0103 \\ 0.0627 & 0.0336 & -0.0688 \\ -0.1025 & -0.0469 & -0.1578 \end{pmatrix} \\ \mathbf{B}_5^4 &= \begin{pmatrix} -0.0005 & 0.0014 & 0.0023 \\ 0.0011 & 0.0003 & -0.0033 \\ 0.0076 & 0.0033 & 0.0115 \end{pmatrix}; & \mathbf{B}_6^4 &= \begin{pmatrix} 0.0011 & -0.0012 & -0.0012 \\ 0.0076 & 0.0034 & 0.0116 \\ -0.0010 & -0.0002 & 0.0036 \end{pmatrix}\end{aligned}$$

Pour la famille de méthodes RSJADE_r, voici la matrice de kurtosis des observations (estimée)

$$\mathcal{C}_{(4)}^{\mathbf{x}} = 1e^5 \begin{pmatrix} 0.9292 & -1.0386 & -0.2371 \\ -1.0386 & 5.7474 & -0.9777 \\ -0.2371 & -0.9777 & 0.4644 \end{pmatrix}, \quad (4.28)$$

qui admet les trois valeurs propres positives suivantes $1e^5$ 6.1092, $1e^5$ 0.9841 et $1e^5$ 0.0476.

4.7 Conclusion

Nous avons proposé dans ce chapitre une famille de méthodes d'ACIM baptisée RSJADE_r, qui exploite uniquement les statistiques d'ordres supérieurs à deux. Cette nouvelle méthode permet, après un blanchiment d'ordre quatre des observations, de résoudre complètement le problème ACIM bruité en bloc diagonalisant conjointement un ensemble de matrices de cumulants contractées d'ordre r ($r > 4$). La famille RSJADE_r a été ensuite comparée favorablement avec les méthodes ACIM à blanchiment d'ordre deux dans le contexte bruité (voir la figure (4.1)). Finalement, afin de mesurer sa performance dans une application réelle, nous avons utilisé cette famille pour séparer les activités électriques fœtales et maternelles à partir d'un mélange de ces activités électriques mesurées via trois électrodes placées sur la peau de la région abdominale de la mère. Nous avons remarqué que toutes les méthodes d'ACIM à blanchiment d'ordre deux séparent ce mélange en estimant quasiment toutes la même matrice de séparation. Un simple coup d'oeil de comparaison pointé vers les figures (4.7) et (4.8) montre que les méthodes à blanchiment d'ordre quatre sont les mieux adaptées à cette application, contrairement à la famille SJADE_r, dont une source ECGM est restée fortement touchée par le cycle cardiaque du fœtus (voir la figure 4.7). Ceci est dû peut-être à une présence de signaux parasites (bruit) dans le mélange ayant bien résisté au blanchiment d'ordre deux et beaucoup moins bien à celui d'ordre quatre.

Les expressions scalaires des cumulants d'un vecteurs aléatoire

Soit $\mathbf{x} = (x_1, \dots, x_n)'$, un vecteur aléatoire réel centré alors, nous avons

$$\text{cum}(x_i, x_j) = \mathbb{E}[x_i x_j] \quad (\text{A.1})$$

$$\text{cum}(x_i, x_j, x_k) = \mathbb{E}[x_i x_j x_k] \quad (\text{A.2})$$

$$\text{cum}(x_i, x_j, x_k, x_l) = \mathbb{E}[x_i x_j x_k x_l] - \mathbb{E}[x_i x_j] \mathbb{E}[x_k x_l] - \mathbb{E}[x_i x_k] \mathbb{E}[x_j x_l] - \mathbb{E}[x_i x_l] \mathbb{E}[x_k x_j] \quad (\text{A.3})$$

$$\begin{aligned} \text{cum}(x_i, x_j, x_k, x_l, x_m) &= \mathbb{E}[x_i x_j x_k x_l x_m] - \mathbb{E}[x_i x_j] \mathbb{E}[x_k x_l x_m] - \mathbb{E}[x_i x_k] \mathbb{E}[x_j x_l x_m] \\ &\quad - \mathbb{E}[x_i x_l] \mathbb{E}[x_k x_j x_m] - \mathbb{E}[x_i x_m] \mathbb{E}[x_k x_l x_j] - \mathbb{E}[x_k x_j] \mathbb{E}[x_i x_l x_m] \\ &\quad - \mathbb{E}[x_l x_j] \mathbb{E}[x_k x_i x_m] - \mathbb{E}[x_m x_j] \mathbb{E}[x_k x_l x_i] - \mathbb{E}[x_k x_l] \mathbb{E}[x_i x_j x_m] \\ &\quad - \mathbb{E}[x_k x_m] \mathbb{E}[x_i x_l x_j] - \mathbb{E}[x_l x_m] \mathbb{E}[x_k x_i x_j] \end{aligned} \quad (\text{A.4})$$

$$\text{cum}(x_i, x_j, x_k, x_l, x_m, x_n) = \mathbb{E}[x_i x_j x_k x_l x_m x_n] - a_{\mathbf{x}} - b_{\mathbf{x}} + 2c_{\mathbf{x}} \quad (\text{A.5})$$

où,

$$\begin{aligned} a_{\mathbf{x}} &= \mathbb{E}(x_i x_j) \mathbb{E}(x_k x_l x_m x_n) + \mathbb{E}(x_i x_k) \mathbb{E}(x_j x_l x_m x_n) + \mathbb{E}(x_i x_l) \mathbb{E}(x_j x_k x_m x_n) + \mathbb{E}(x_i x_m) \mathbb{E}(x_j x_k x_l x_n) \\ &\quad + \mathbb{E}(x_i x_n) \mathbb{E}(x_j x_k x_l x_m) + \mathbb{E}(x_j x_k) \mathbb{E}(x_i x_l x_m x_n) + \mathbb{E}(x_j x_l) \mathbb{E}(x_i x_k x_m x_n) + \mathbb{E}(x_j x_m) \mathbb{E}(x_i x_k x_l x_n) \\ &\quad + \mathbb{E}(x_j x_n) \mathbb{E}(x_i x_k x_l x_m) + \mathbb{E}(x_k x_l) \mathbb{E}(x_i x_j x_m x_n) + \mathbb{E}(x_k x_m) \mathbb{E}(x_i x_j x_l x_n) + \mathbb{E}(x_k x_n) \mathbb{E}(x_i x_j x_l x_m) \\ &\quad + \mathbb{E}(x_l x_m) \mathbb{E}(x_i x_j x_k x_n) + \mathbb{E}(x_l x_n) \mathbb{E}(x_i x_j x_k x_m) + \mathbb{E}(x_m x_n) \mathbb{E}(x_i x_j x_k x_l) \end{aligned} \quad (\text{A.6})$$

$$\begin{aligned} b_{\mathbf{x}} &= \mathbb{E}(x_i x_j x_k) \mathbb{E}(x_l x_m x_n) + \mathbb{E}(x_i x_j x_l) \mathbb{E}(x_k x_m x_n) + \mathbb{E}(x_i x_j x_m) \mathbb{E}(x_k x_l x_n) + \mathbb{E}(x_i x_j x_n) \mathbb{E}(x_k x_l x_m) \\ &\quad + \mathbb{E}(x_i x_k x_l) \mathbb{E}(x_j x_m x_n) + \mathbb{E}(x_i x_k x_m) \mathbb{E}(x_j x_l x_n) + \mathbb{E}(x_i x_k x_n) \mathbb{E}(x_j x_l x_m) + \mathbb{E}(x_i x_l x_m) \mathbb{E}(x_j x_k x_n) \\ &\quad + \mathbb{E}(x_i x_l x_n) \mathbb{E}(x_j x_k x_m) + \mathbb{E}(x_i x_m x_n) \mathbb{E}(x_j x_k x_l) \end{aligned} \quad (\text{A.7})$$

$$\begin{aligned} c_{\mathbf{x}} &= \mathbb{E}(x_i x_j) \mathbb{E}(x_k x_l) \mathbb{E}(x_m x_n) + \mathbb{E}(x_i x_j) \mathbb{E}(x_k x_m) \mathbb{E}(x_l x_n) + \mathbb{E}(x_i x_j) \mathbb{E}(x_k x_n) \mathbb{E}(x_l x_m) \\ &\quad + \mathbb{E}(x_i x_k) \mathbb{E}(x_j x_l) \mathbb{E}(x_m x_n) + \mathbb{E}(x_i x_k) \mathbb{E}(x_j x_m) \mathbb{E}(x_l x_n) + \mathbb{E}(x_i x_k) \mathbb{E}(x_j x_n) \mathbb{E}(x_l x_m) \\ &\quad + \mathbb{E}(x_i x_l) \mathbb{E}(x_j x_k) \mathbb{E}(x_m x_n) + \mathbb{E}(x_i x_l) \mathbb{E}(x_j x_m) \mathbb{E}(x_k x_n) + \mathbb{E}(x_i x_l) \mathbb{E}(x_j x_n) \mathbb{E}(x_k x_m) \\ &\quad + \mathbb{E}(x_i x_m) \mathbb{E}(x_j x_k) \mathbb{E}(x_l x_n) + \mathbb{E}(x_i x_m) \mathbb{E}(x_j x_l) \mathbb{E}(x_k x_n) + \mathbb{E}(x_i x_m) \mathbb{E}(x_j x_n) \mathbb{E}(x_k x_l) \\ &\quad + \mathbb{E}(x_i x_n) \mathbb{E}(x_j x_k) \mathbb{E}(x_l x_m) + \mathbb{E}(x_i x_n) \mathbb{E}(x_j x_l) \mathbb{E}(x_k x_m) + \mathbb{E}(x_i x_n) \mathbb{E}(x_j x_m) \mathbb{E}(x_k x_l) \end{aligned} \quad (\text{A.8})$$

Produit de Kronecker et produit étoile

Définition 12 (matrice partitionnée). Une matrice $\mathbf{A} \in \mathbb{R}^{p \times q}$ est dite matrice partitionnée (ou matrice par blocs), si elle est constituée par les matrices $\mathbf{A}_{i,j} \in \mathbb{R}^{p_i \times q_j}$, de la façon suivante :

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{1,1} & \cdots & \mathbf{A}_{1,m} \\ \vdots & \ddots & \vdots \\ \mathbf{A}_{n,1} & \cdots & \mathbf{A}_{n,m} \end{pmatrix}, \quad p = \sum_{i=1}^n p_i; \quad q = \sum_{j=1}^m q_j \quad (\text{B.1})$$

B.1 Produit de Kronecker

Définition 13 (Produit de Kronecker). Le produit de Kronecker de deux matrices \mathbf{A} et \mathbf{B} est la matrice notée $\mathbf{A} \otimes \mathbf{B}$ définie par ses partitions

$$(\mathbf{A} \otimes \mathbf{B})_{i,j} = a_{i,j} \mathbf{B} \quad (\text{B.2})$$

Nous rappelons quelques propriétés de ce produit.

Proposition 36. Soient $\mathbf{A} \in \mathbb{R}^{p \times q}$, $\mathbf{B} \in \mathbb{R}^{q \times n}$, $\mathbf{C} \in \mathbb{R}^{r \times t}$ et $\mathbf{D} \in \mathbb{R}^{t \times s}$. On a les égalités suivantes :

$$(\mathbf{A} \otimes \mathbf{B})' = (\mathbf{A}' \otimes \mathbf{B}') \quad (\text{B.3})$$

$$(\mathbf{A} + \mathbf{B}) \otimes (\mathbf{C} + \mathbf{D}) = \mathbf{A} \otimes \mathbf{C} + \mathbf{A} \otimes \mathbf{D} + \mathbf{B} \otimes \mathbf{C} + \mathbf{B} \otimes \mathbf{D} \quad (\text{B.4})$$

$$\mathbf{AB} \otimes \mathbf{CD} = (\mathbf{A} \otimes \mathbf{C})(\mathbf{B} \otimes \mathbf{D}) \quad (\text{B.5})$$

$$(\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C}) \quad (\text{B.6})$$

Si \mathbf{A} et \mathbf{B} sont inversibles, alors,

$$(\mathbf{A} \otimes \mathbf{B})^{-1} = (\mathbf{A}^{-1} \otimes \mathbf{B}^{-1}) \quad (\text{B.7})$$

B.2 Matrice de commutation

Soit $\mathbf{A} \in \mathbb{R}^{m \times n}$. Lorsqu'on a besoin d'exprimer $\text{vec} \mathbf{A}$ en fonction de $\text{vec} \mathbf{A}'$, la matrice partitionnée dite de commutation permet ce passage. Voici sa définition :

Définition 14 ([Magnus 1979]). On appelle $\mathbf{K}_{mn} \in \mathbb{R}^{mn \times mn}$, la matrice de commutation définie par

$$\mathbf{K}_{mn} = \sum_{i=1}^n \sum_{j=1}^m \mathbf{a}_i \mathbf{b}_j' \otimes \mathbf{b}_j \mathbf{a}_i' \quad (\text{B.8})$$

où \mathbf{a}_i est une base de \mathbb{R}^m et \mathbf{b}_i celle de \mathbb{R}^n . Cette matrice est en plus une matrice de permutation.

Exemple 1. Si $m = 3$ et $n = 2$, alors, \mathbf{K}_{32} est la matrice 6×6 suivante :

$$\mathbf{K}_{32} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (\text{B.9})$$

Nous rappelons dans la proposition suivante quelques propriétés intéressantes des matrices de commutation

Proposition 37.

$$\mathbf{K}_{nm} \text{vec} \mathbf{A} = \text{vec} \mathbf{A}' \quad (\text{B.10})$$

$$\mathbf{K}'_{mn} = \mathbf{K}_{nm} \quad (\text{B.11})$$

$$\mathbf{K}_{nm} \mathbf{K}_{mn} = \mathbf{I}_{mn} \quad (\text{B.12})$$

$$\mathbf{K}_{mrs} = (\mathbf{I}_r \otimes \mathbf{K}_{ms}) (\mathbf{K}_{mr} \otimes \mathbf{I}_s) \quad (\text{B.13})$$

Il est possible de commuter le produit de Kronecker à l'aide des matrices des commutations.

Proposition 38. Soient $\mathbf{A} \in \mathbb{R}^{p \times q}$ et $\mathbf{B} \in \mathbb{R}^{r \times s}$. Nous avons

$$\mathbf{A} \otimes \mathbf{B} = \mathbf{K}_{rp} (\mathbf{B} \otimes \mathbf{A}) \mathbf{K}_{qs} \quad (\text{B.14})$$

$$\text{vec}(\mathbf{A} \otimes \mathbf{B}) = (\mathbf{I}_q \otimes \mathbf{K}_{ps} \otimes \mathbf{I}_r) \text{vec} \mathbf{A} \otimes \text{vec} \mathbf{B} \quad (\text{B.15})$$

Proposition 39. Soient $\mathbf{A} \in \mathbb{R}^{p \times q}$, $\mathbf{B} \in \mathbb{R}^{q \times r}$ et $\mathbf{C} \in \mathbb{R}^{r \times s}$. Alors,

$$\text{vec}(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A}) \text{vec} \mathbf{B} \quad (\text{B.16})$$

B.3 Le produit étoile

Définition 15. Soient \mathbf{A} une matrice de taille $p \times q$ et \mathbf{B} une matrice de taille $rp \times sq$. Le produit étoile de \mathbf{A} et \mathbf{B} , noté $\mathbf{A} * \mathbf{B}$, est la matrice de taille $r \times s$ définie par

$$\mathbf{A} * \mathbf{B} = \sum_{i_1, i_2=1}^{p, q} a_{i_1, i_2} \mathbf{B}_{i_1, i_2} \quad (\text{B.17})$$

où \mathbf{B}_{i_1, i_2} est le bloc indexé par (i_1, i_2) dans le partage de la matrice \mathbf{B} en matrices de taille $r \times s$.

Donnons quelques propriétés intéressantes du produit étoile.

Proposition 40. (P1) Soient \mathbf{A} et \mathbf{B} deux matrices de même taille. Alors

$$\mathbf{A} * \mathbf{B} = \text{Tr}(\mathbf{A}' \mathbf{B}). \quad (\text{B.18})$$

(P2) Soient \mathbf{A} , \mathbf{B} et \mathbf{C} trois matrices de tailles respectives $p \times q$, $pr \times sq$ et $m \times n$. Alors on a :

$$\mathbf{A} * (\mathbf{B} \otimes \mathbf{C}) = (\mathbf{A} * \mathbf{B}) \otimes \mathbf{C} \quad (\text{B.19})$$

(P3) Soient \mathbf{A} , \mathbf{B} et \mathbf{C} trois matrices de tailles respectives $p \times q$, $q \times r$ et $r \times s$. Alors on a :

$$\mathbf{ABC} = \mathbf{B} * \text{vec} \mathbf{A} \text{vec}' \mathbf{C}' = \mathbf{B}' * (\mathbf{C} \otimes \mathbf{I}_p) \mathbf{K}_{ps} (\mathbf{A} \otimes \mathbf{I}_s), \quad (\text{B.20})$$

(P4) Soient \mathbf{a} et \mathbf{b} deux vecteurs colonnes de dimensions respectives p et q . Soit \mathbf{M} une matrice de taille $pr \times qs$. Alors on a :

$$(\mathbf{a}' \otimes \mathbf{I}_r) \mathbf{M} (\mathbf{b} \otimes \mathbf{I}_s) = \mathbf{ab}' * \mathbf{M} \quad (\text{B.21})$$

Démonstration. Les propriétés (P1), (P2) et (P3) sont données sans démonstration dans [MacRae 1974]; pour une démonstration, voir [Kollo 1991, pages. 34-37]. Nous allons montrer uniquement (P4) qui sera

utilisé à la fin de la preuve du théorème 22. La matrice \mathbf{M} , de taille $pr \times qs$, peut être partagée en blocs de taille $r \times s$ de la manière suivante

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_{1,1} & \mathbf{M}_{1,2} & \cdots & \mathbf{M}_{1,q} \\ \mathbf{M}_{2,1} & \mathbf{M}_{2,2} & \cdots & \mathbf{M}_{2,q} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{M}_{p,1} & \mathbf{M}_{p,2} & \cdots & \mathbf{M}_{p,q} \end{pmatrix} \quad (\text{B.22})$$

Si $\mathbf{a} = (a_1, \dots, a_p)'$ et $\mathbf{b} = (b_1, \dots, b_q)'$, nous avons

$$\begin{aligned} (\mathbf{a}' \otimes \mathbf{I}_r) \mathbf{M} (\mathbf{b} \otimes \mathbf{I}_s) &= \begin{pmatrix} a_1 \mathbf{I}_r & a_2 \mathbf{I}_r & \cdots & a_p \mathbf{I}_r \end{pmatrix} \begin{pmatrix} \mathbf{M}_{1,1} & \mathbf{M}_{1,2} & \cdots & \mathbf{M}_{1,q} \\ \mathbf{M}_{2,1} & \mathbf{M}_{2,2} & \cdots & \mathbf{M}_{2,q} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{M}_{p,1} & \mathbf{M}_{p,2} & \cdots & \mathbf{M}_{p,q} \end{pmatrix} \begin{pmatrix} b_1 \mathbf{I}_s \\ b_2 \mathbf{I}_s \\ \vdots \\ b_q \mathbf{I}_s \end{pmatrix} \\ &= \begin{pmatrix} \sum_{i=1}^p a_i \mathbf{M}_{i,1} & \sum_{i=1}^p a_i \mathbf{M}_{i,2} & \cdots & \sum_{i=1}^p a_i \mathbf{M}_{i,p} \end{pmatrix} \begin{pmatrix} b_1 \mathbf{I}_s \\ b_2 \mathbf{I}_s \\ \vdots \\ b_q \mathbf{I}_s \end{pmatrix} \\ &= b_1 \sum_{i=1}^p a_i \mathbf{M}_{i,1} + b_2 \sum_{i=1}^p a_i \mathbf{M}_{i,2} + \cdots + b_q \sum_{i=1}^p a_i \mathbf{M}_{i,p} = \sum_{i=1}^p \sum_{j=1}^q a_i b_j \mathbf{M}_{i,j} \\ &= \mathbf{ab}' * \mathbf{M} \end{aligned} \quad (\text{B.23})$$

□

Sur la réduction de l'ensemble de matrices des cumulants contractées à bloc diagonaliser conjointement

Dans le but de réduire le nombre de matrices à bloc diagonaliser conjointement, nous allons, maintenant donner en fonction de la base canonique de \mathbb{R}^n , une base que nous notons $\mathcal{E}_{i_1, i_2, \dots, i_{r-2}}$ orthogonale dans le sens suivant :

$$\begin{cases} \mathbf{vec}' \mathcal{E}_{i_1, i_2, \dots, i_r} \mathbf{vec} \mathcal{E}_{j_1, j_2, \dots, j_{r-2}} = 0 & \text{si } (i_1, i_2, \dots, i_{r-2}) \neq (j_1, j_2, \dots, j_{r-2}) \\ \mathbf{vec}' \mathcal{E}_{i_1, i_2, \dots, i_{r-2}} \mathbf{vec} \mathcal{E}_{i_1, i_2, \dots, i_{r-2}} = 1 \end{cases} \quad (\text{C.1})$$

qui permette d'engendrer l'ensemble $\mathcal{M}^{(r)}$ (3.54) pour les ordres 3 à 6 (les seuls utilisés dans ce document).

— Ordre 3

$$\mathcal{M}^{(3)} = \left\{ \mathcal{Q}_{(3)}^{\mathbf{x}}[\mathcal{E}_i], \quad 1 \leq i \leq n \right\} \quad (\text{C.2})$$

où $\mathcal{E}_i = \mathbf{e}_i$.

— Ordre 4.

$$\mathcal{M}^{(4)} = \left\{ \mathcal{Q}_{(4)}^{\mathbf{x}}[\mathcal{E}_{i,j}], \quad 1 \leq i \leq j \leq n \right\} \quad (\text{C.3})$$

où

$$\mathcal{E}_{i,j} = \begin{cases} \mathbf{e}_i \mathbf{e}'_i & \text{si } i = j \\ \frac{1}{\sqrt{2}} (\mathbf{e}_i \mathbf{e}'_j + \mathbf{e}_j \mathbf{e}'_i) & \text{si } i < j \end{cases} \quad (\text{C.4})$$

Cette base a été donner par [Cardoso 1999].

— Ordre 5.

$$\mathcal{M}^{(5)} = \left\{ \mathcal{Q}_{(5)}^{\mathbf{x}}[\mathcal{E}_{i,j,k}], \quad 1 \leq i \leq j \leq k \leq n \right\} \quad (\text{C.5})$$

où

$$\mathcal{E}_{i,j,k} = \begin{cases} \mathbf{e}_i \otimes \mathbf{e}_i \mathbf{e}'_i & \text{si } i = j = k \\ \frac{1}{\sqrt{3}} (\mathbf{e}_k \otimes \mathbf{e}_i \mathbf{e}'_i + \mathbf{e}_i \otimes \mathbf{e}_k \mathbf{e}'_i + \mathbf{e}_i \otimes \mathbf{e}_i \mathbf{e}'_k) & \text{si } i = j < k \\ \frac{1}{\sqrt{3}} (\mathbf{e}_i \otimes \mathbf{e}_k \mathbf{e}'_k + \mathbf{e}_k \otimes \mathbf{e}_i \mathbf{e}'_k + \mathbf{e}_k \otimes \mathbf{e}_k \mathbf{e}'_i) & \text{si } i < j = k \\ \frac{1}{\sqrt{6}} (\mathbf{e}_i \otimes \mathbf{e}_j \mathbf{e}'_k + \mathbf{e}_i \otimes \mathbf{e}_k \mathbf{e}'_j + \mathbf{e}_j \otimes \mathbf{e}_i \mathbf{e}'_k + \mathbf{e}_j \otimes \mathbf{e}_k \mathbf{e}'_i + \mathbf{e}_k \otimes \mathbf{e}_i \mathbf{e}'_j + \mathbf{e}_k \otimes \mathbf{e}_j \mathbf{e}'_i) & \text{si } i < j < k \end{cases} \quad (\text{C.6})$$

— Order 6.

$$\mathcal{M}^{(6)} = \left\{ \mathcal{Q}_{(6)}^{\mathbf{x}}[\mathcal{E}_{i,j,k,l}], \quad 1 \leq i \leq j \leq k \leq l \leq n \right\} \quad (\text{C.7})$$

où

$$\mathcal{E}_{i,j,k,l} = \begin{cases} \mathbf{e}_i \mathbf{e}'_i \otimes \mathbf{e}_i \mathbf{e}'_i & \text{si } i = j = k = l ; \\ \frac{1}{2} (\mathbf{e}_i \mathbf{e}'_i \otimes \mathbf{e}_i \mathbf{e}'_l + \mathbf{e}_i \mathbf{e}'_i \otimes \mathbf{e}_l \mathbf{e}'_i + \mathbf{e}_i \mathbf{e}'_l \otimes \mathbf{e}_i \mathbf{e}'_i + \mathbf{e}_l \mathbf{e}'_i \otimes \mathbf{e}_i \mathbf{e}'_i) & \text{si } i = j = k < l ; \\ \frac{1}{2} (\mathbf{e}_l \mathbf{e}'_l \otimes \mathbf{e}_l \mathbf{e}'_i + \mathbf{e}_l \mathbf{e}'_l \otimes \mathbf{e}_i \mathbf{e}'_l + \mathbf{e}_l \mathbf{e}'_i \otimes \mathbf{e}_l \mathbf{e}'_l + \mathbf{e}_i \mathbf{e}'_l \otimes \mathbf{e}_l \mathbf{e}'_l) & \text{si } i < j = k = l ; \\ \frac{1}{\sqrt{6}} (\mathbf{e}_i \mathbf{e}'_i \otimes \mathbf{e}_l \mathbf{e}'_l + \mathbf{e}_i \mathbf{e}'_l \otimes \mathbf{e}_i \mathbf{e}'_l + \mathbf{e}_i \mathbf{e}'_l \otimes \mathbf{e}_l \mathbf{e}'_i + \mathbf{e}_l \mathbf{e}'_i \otimes \mathbf{e}_i \mathbf{e}'_l + \mathbf{e}_l \mathbf{e}'_i \otimes \mathbf{e}_l \mathbf{e}'_i + \mathbf{e}_l \mathbf{e}'_l \otimes \mathbf{e}_i \mathbf{e}'_i) & \text{si } i = j < k = l ; \\ \frac{1}{\sqrt{12}} \left(\mathbf{e}_i \mathbf{e}'_i \otimes \mathbf{e}_k \mathbf{e}'_l + \mathbf{e}_i \mathbf{e}'_i \otimes \mathbf{e}_l \mathbf{e}'_k + \mathbf{e}_i \mathbf{e}'_k \otimes \mathbf{e}_i \mathbf{e}'_l + \mathbf{e}_i \mathbf{e}'_k \otimes \mathbf{e}_l \mathbf{e}'_i + \mathbf{e}_i \mathbf{e}'_l \otimes \mathbf{e}_i \mathbf{e}'_k + \mathbf{e}_i \mathbf{e}'_l \otimes \mathbf{e}_k \mathbf{e}'_i \right. \\ \left. + \mathbf{e}_k \mathbf{e}'_i \otimes \mathbf{e}_i \mathbf{e}'_l + \mathbf{e}_k \mathbf{e}'_i \otimes \mathbf{e}_l \mathbf{e}'_i + \mathbf{e}_l \mathbf{e}'_i \otimes \mathbf{e}_i \mathbf{e}'_k + \mathbf{e}_l \mathbf{e}'_i \otimes \mathbf{e}_k \mathbf{e}'_i + \mathbf{e}_k \mathbf{e}'_l \otimes \mathbf{e}_i \mathbf{e}'_i + \mathbf{e}_l \mathbf{e}'_k \otimes \mathbf{e}_i \mathbf{e}'_i \right) & \text{si } i = j < k < l ; \\ \frac{1}{\sqrt{12}} \left(\mathbf{e}_j \mathbf{e}'_j \otimes \mathbf{e}_i \mathbf{e}'_l + \mathbf{e}_j \mathbf{e}'_j \otimes \mathbf{e}_l \mathbf{e}'_i + \mathbf{e}_j \mathbf{e}'_i \otimes \mathbf{e}_j \mathbf{e}'_l + \mathbf{e}_j \mathbf{e}'_i \otimes \mathbf{e}_l \mathbf{e}'_j + \mathbf{e}_j \mathbf{e}'_l \otimes \mathbf{e}_j \mathbf{e}'_i + \mathbf{e}_j \mathbf{e}'_l \otimes \mathbf{e}_i \mathbf{e}'_j \right. \\ \left. + \mathbf{e}_i \mathbf{e}'_j \otimes \mathbf{e}_j \mathbf{e}'_l + \mathbf{e}_i \mathbf{e}'_j \otimes \mathbf{e}_l \mathbf{e}'_j + \mathbf{e}_l \mathbf{e}'_j \otimes \mathbf{e}_j \mathbf{e}'_i + \mathbf{e}_l \mathbf{e}'_j \otimes \mathbf{e}_i \mathbf{e}'_j + \mathbf{e}_i \mathbf{e}'_l \otimes \mathbf{e}_j \mathbf{e}'_j + \mathbf{e}_l \mathbf{e}'_i \otimes \mathbf{e}_j \mathbf{e}'_j \right) & \text{si } i < j = k < l ; \\ \frac{1}{\sqrt{12}} \left(\mathbf{e}_l \mathbf{e}'_l \otimes \mathbf{e}_j \mathbf{e}'_i + \mathbf{e}_l \mathbf{e}'_l \otimes \mathbf{e}_i \mathbf{e}'_j + \mathbf{e}_l \mathbf{e}'_j \otimes \mathbf{e}_l \mathbf{e}'_i + \mathbf{e}_l \mathbf{e}'_j \otimes \mathbf{e}_i \mathbf{e}'_l + \mathbf{e}_l \mathbf{e}'_i \otimes \mathbf{e}_l \mathbf{e}'_j + \mathbf{e}_l \mathbf{e}'_i \otimes \mathbf{e}_j \mathbf{e}'_l \right. \\ \left. + \mathbf{e}_j \mathbf{e}'_l \otimes \mathbf{e}_l \mathbf{e}'_i + \mathbf{e}_j \mathbf{e}'_l \otimes \mathbf{e}_i \mathbf{e}'_l + \mathbf{e}_i \mathbf{e}'_l \otimes \mathbf{e}_l \mathbf{e}'_j + \mathbf{e}_i \mathbf{e}'_l \otimes \mathbf{e}_j \mathbf{e}'_l + \mathbf{e}_j \mathbf{e}'_i \otimes \mathbf{e}_l \mathbf{e}'_l + \mathbf{e}_i \mathbf{e}'_j \otimes \mathbf{e}_l \mathbf{e}'_l \right) & \text{si } i < j < k = l ; \\ \frac{1}{\sqrt{24}} \left(\mathbf{e}_i \mathbf{e}'_j \otimes \mathbf{e}_k \mathbf{e}'_l + \mathbf{e}_i \mathbf{e}'_j \otimes \mathbf{e}_l \mathbf{e}'_k + \mathbf{e}_i \mathbf{e}'_k \otimes \mathbf{e}_j \mathbf{e}'_l + \mathbf{e}_i \mathbf{e}'_k \otimes \mathbf{e}_l \mathbf{e}'_j + \mathbf{e}_i \mathbf{e}'_l \otimes \mathbf{e}_j \mathbf{e}'_k + \mathbf{e}_i \mathbf{e}'_l \otimes \mathbf{e}_k \mathbf{e}'_j \right. \\ \left. + \mathbf{e}_j \mathbf{e}'_i \otimes \mathbf{e}_k \mathbf{e}'_l + \mathbf{e}_j \mathbf{e}'_i \otimes \mathbf{e}_l \mathbf{e}'_k + \mathbf{e}_j \mathbf{e}'_k \otimes \mathbf{e}_i \mathbf{e}'_l + \mathbf{e}_j \mathbf{e}'_k \otimes \mathbf{e}_l \mathbf{e}'_i + \mathbf{e}_j \mathbf{e}'_l \otimes \mathbf{e}_i \mathbf{e}'_k + \mathbf{e}_j \mathbf{e}'_l \otimes \mathbf{e}_k \mathbf{e}'_i \right. \\ \left. + \mathbf{e}_k \mathbf{e}'_i \otimes \mathbf{e}_j \mathbf{e}'_l + \mathbf{e}_k \mathbf{e}'_i \otimes \mathbf{e}_l \mathbf{e}'_j + \mathbf{e}_k \mathbf{e}'_j \otimes \mathbf{e}_i \mathbf{e}'_l + \mathbf{e}_k \mathbf{e}'_j \otimes \mathbf{e}_l \mathbf{e}'_i + \mathbf{e}_k \mathbf{e}'_l \otimes \mathbf{e}_j \mathbf{e}'_i + \mathbf{e}_k \mathbf{e}'_l \otimes \mathbf{e}_i \mathbf{e}'_j \right. \\ \left. + \mathbf{e}_l \mathbf{e}'_i \otimes \mathbf{e}_j \mathbf{e}'_k + \mathbf{e}_l \mathbf{e}'_i \otimes \mathbf{e}_k \mathbf{e}'_j + \mathbf{e}_l \mathbf{e}'_j \otimes \mathbf{e}_i \mathbf{e}'_k + \mathbf{e}_l \mathbf{e}'_j \otimes \mathbf{e}_k \mathbf{e}'_i + \mathbf{e}_l \mathbf{e}'_k \otimes \mathbf{e}_i \mathbf{e}'_j + \mathbf{e}_l \mathbf{e}'_k \otimes \mathbf{e}_j \mathbf{e}'_i \right) & \text{si } i < j < k < l ; \end{cases} \quad (\text{C.8})$$

Bibliographie

- [A. Belouchrani 1997] J. F. Cardoso A. Belouchrani K. Abed-Meraim et E. Moulines. *A blind source separation technique using second order statistics*. IEEE Trans. Signal Process, vol. 45, no. 2, pages 434–444, Feb 1997. (Cité en pages 9, 10, 60 et 67.)
- [Abadir 2005] K.M. Abadir et J.R. Magnus. Matrix algebra. Econometric Exercises. Cambridge University Press, 2005. (Cité en page 54.)
- [Abed-Meraim 2004] Karim Abed-Meraim et Adel Belouchrani. *Algorithms for joint block diagonalization*. In Proc. EUSIPCO, volume 4, pages 209–212, 2004. (Cité en pages 15, 20, 50 et 57.)
- [Akaho 1999] S. Akaho, Y. Kiuchi et S. Umeyama. *MICA : multimodal independent component analysis*. In Neural Networks, 1999. IJCNN '99. International Joint Conference on, volume 2, pages 927–932 vol.2, 1999. (Cité en page 19.)
- [Albera 2002] L. Albera et P. Comon. *Asymptotic performance of contrast-based blind source separation algorithms*. In Second IEEE Sensor Array and Multichannel Signal Processing Workshop, Rosslyn, VA, August 2002. (Cité en pages 11, 26 et 41.)
- [Albera 2003] L. Albera, A. Ferreol, P. Comon et P. Chevalier. *Sixth order blind identification of underdetermined mixtures (BIRTH) of sources*. In 4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA2003), Nara, Japan, pages 909–914, April 2003. (Cité en pages 10, 23, 41 et 42.)
- [Albera 2004] L. Albera, A. Ferreol, P. Comon et P. Chevalier. *Blind Identification of Over-complete Mixtures of sources (BIOME)*. Linear Algebra and its Applications, vol. 391, pages 1–30, November 2004. (Cité en pages 10, 23, 37, 41, 67 et 72.)
- [Albera 2005] L. Albera, A. Ferreol, P. Chevalier et P. Comon. *ICAR : a tool for blind source separation using fourth-order statistics only*. Signal Processing, IEEE Transactions on, vol. 53, no. 10, pages 3633–3643, 2005. (Cité en pages 10, 23, 41, 42, 67, 72, 73 et 74.)
- [Almeida 2009] Miguel Almeida et Ricardo Vigário. *Source Separation of Phase-Locked Subspaces*. In Proceedings of the 8th International Conference on Independent Component Analysis and Signal Separation, ICA '09, pages 203–210, Berlin, Heidelberg, 2009. Springer-Verlag. (Cité en page 19.)
- [Almeida 2011] Miguel Almeida, J-H Schleimer, José Mario Bioucas-Dias et Ricardo Vigário. *Source separation and clustering of phase-locked subspaces*. Neural Networks, IEEE Transactions on, vol. 22, no. 9, pages 1419–1434, 2011. (Cité en page 19.)
- [Amari 1996] S. Amari, A. Cichocki et H. H. Yang. *A New Learning Algorithm for Blind Signal Separation*. In Advances in Neural Information Processing Systems, pages 757–763. MIT Press, 1996. (Cité en pages 8 et 59.)
- [Antoni 2005] J Antoni, François Guillet, Mohamed El Badaoui et Frédéric Bonnardot. *Blind separation of convolved cyclostationary processes*. Signal processing, vol. 85, no. 1, pages 51–66, 2005. (Cité en page 10.)
- [Azzerboni 2005] Bruno Azzerboni, Fabio La Foresta, Nadia Mammone et Francesco Carlo Morabito. *A new approach based on wavelet-ICA algorithms for fetal electrocardiogram extraction*. In ESANN, pages 193–198, 2005. (Cité en page 76.)
- [Bach 2003a] Francis R. Bach et Michael I. Jordan. *Beyond independent components : trees and clusters*. J. Mach. Learn. Res., vol. 4, pages 1205–1233, Décembre 2003. (Cité en pages 13 et 14.)
- [Bach 2003b] Francis R. Bach et Michael I. Jordan. *Finding Clusters In Independent Component Analysis*. IN : 4TH INTL. SYMP. ON INDEPENDENT COMPONENT ANALYSIS AND SIGNAL SEPARATION (ICA2003, pages 891–896, 2003. (Cité en page 14.)

- [Bach 2003c] Francis R. Bach et Michael I. Jordan. *Kernel independent component analysis*. J. Mach. Learn. Res., vol. 3, pages 1–48, Mars 2003. (Cit  en page 14.)
- [Bell 1995] Anthony J. Bell et Terrence J. Sejnowski. *An information-maximization approach to blind separation and blind deconvolution*. Neural Comput., vol. 7, no. 6, pages 1129–1159, Novembre 1995. (Cit  en pages 8 et 67.)
- [Belouchrani 1997] Adel Belouchrani, Moeness G Amin et Karim Abed-Meraim. *Direction finding in correlated noise fields based on joint block-diagonalization of spatio-temporal correlation matrices*. Signal Processing Letters, IEEE, vol. 4, no. 9, pages 266–268, 1997. (Cit  en page 19.)
- [Belouchrani 2000] Adel Belouchrani et Andrzej Cichocki. *A Robust Whitening Procedure in Blind Source Separation Context*. Electronics Letters, vol. 36, pages 2050–2051, 2000. (Cit  en pages 9, 10 et 67.)
- [Belouchrani 2001] A Belouchrani, K Abed-Meraim, MG Amin et AM Zoubir. *Joint anti-diagonalization for blind source separation*. In Acoustics, Speech, and Signal Processing, 2001. Proceedings.(ICASSP’01). 2001 IEEE International Conference on, volume 5, pages 2789–2792. IEEE, 2001. (Cit  en page 19.)
- [Blanchard 2006] Gilles Blanchard, Motoaki Kawanabe, Masashi Sugiyama, Vladimir Spokoiny et Klaus-Robert M ller. *In search of non-Gaussian components of a high-dimensional distribution*. The Journal of Machine Learning Research, vol. 7, pages 247–282, 2006. (Cit  en pages 19 et 67.)
- [Bonhomme 2008] St phane Bonhomme et Jean-Marc Robin. *Consistent noisy independent component analysis*. CeMMAP working papers CWP04/08, Centre for Microdata Methods and Practice, Institute for Fiscal Studies, F vrier 2008. (Cit  en pages 10 et 67.)
- [Bousbia-Salah 2001] H. Bousbia-Salah, A. Belouchrani et K. Abed-Meraim. *Blind separation of non stationary sources using joint block diagonalization*. In Statistical Signal Processing, 2001. Proceedings of the 11th IEEE Signal Processing Workshop on, pages 448–451, 2001. (Cit  en pages 19 et 20.)
- [Boustany 2005] Roger Boustany et Jerome Antoni. *A subspace method for the blind extraction of a cyclostationary source : Application to rolling element bearing diagnostics*. Mechanical Systems and Signal Processing, vol. 19, no. 6, pages 1245–1259, 2005. (Cit  en pages 10 et 76.)
- [Boustany 2008] Roger Boustany et J r me Antoni. *Blind extraction of a cyclostationary signal using reduced-rank cyclic regression ?a unifying approach*. Mechanical Systems and Signal Processing, vol. 22, no. 3, pages 520–541, 2008. (Cit  en page 10.)
- [Bradaric 2000] Ivan Bradaric, Athina P. Petropulu et Konstantinos I. Diamantaras. *On Resolving the Column Permutation Ambiguity in the Estimates of MIMO System Response*, 2000. (Cit  en pages 7 et 8.)
- [Camargo-Olivares 2011] Jos  Luis Camargo-Olivares, Rub n Mart n-Clemente, Susana Hornillo-Mellado, MM Elena et Isabel Rom n. *The maternal abdominal ECG as input to MICA in the fetal ECG extraction problem*. Signal Processing Letters, IEEE, vol. 18, no. 3, pages 161–164, 2011. (Cit  en page 76.)
- [Cardoso 1989a] J.-F. Cardoso. *Blind Identification Of Independent Components With Higher-order Statistics*. In Higher-Order Spectral Analysis, 1989. Workshop on, pages 157–162, 1989. (Cit  en page 8.)
- [Cardoso 1989b] J.-F. Cardoso. *Source separation using higher order moments*. In Acoustics, Speech, and Signal Processing, 1989. ICASSP-89., 1989 International Conference on, pages 2109–2112 vol.4, 1989. (Cit  en pages 8, 9 et 68.)
- [Cardoso 1991] J. Cardoso. *Super-symmetric decomposition of the fourth-order cumulant tensor. Blind identification of more sources than sensors*. In Acoustics, Speech, and Signal Processing, 1991. ICASSP-91., 1991 International Conference on, pages 3109–3112 vol.5, 1991. (Cit  en page 10.)

- [Cardoso 1993] J.-F. Cardoso et A. Soulomiac. *Blind beamforming for non gaussian signals*. IEE-Proceedings-F, vol. 140, no. 6, pages 362–370, 1993. (Cité en pages vii, 8, 9, 15, 19, 20, 23, 39, 44, 50, 57, 59, 60, 67 et 76.)
- [Cardoso 1994] Jean-François Cardoso et al. *On the performance of orthogonal source separation algorithms*. In Proc. EUSIPCO, volume 94, pages 776–779. Edinburgh, UK, 1994. (Cité en page 10.)
- [Cardoso 1996] Jean-François Cardoso et Antoine Souloumiac. *Jacobi Angles for Simultaneous Diagonalization*. SIAM J. Matrix Anal. Appl., vol. 17, no. 1, pages 161–164, Janvier 1996. (Cité en pages 8, 19 et 53.)
- [Cardoso 1998] J. Cardoso. *Multidimensional independent component analysis*. In Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on, volume 4, pages 1941–1944 vol.4, 1998. (Cité en pages 11, 13, 14, 15, 41, 57 et 76.)
- [Cardoso 1999] Jean-François Cardoso. *High-order contrasts for independent component analysis*. Neural computation, vol. 11, no. 1, pages 157–192, 1999. (Cité en pages 45 et 89.)
- [Casey 2000] Michael A Casey et Alex Westner. *Separation of mixed audio sources by independent subspace analysis*. In Proceedings of the International Computer Music Conference, pages 154–161, 2000. (Cité en page 13.)
- [Castella 2007] Marc Castella et Pierre Comon. *Blind separation of instantaneous mixtures of dependent sources*. In Independent Component Analysis and Signal Separation, pages 9–16. Springer, 2007. (Cité en pages 15 et 39.)
- [Chabriel 2008] Gilles Chabriel, Jean Barrère, Nadège Thirion-Moreau et Eric Moreau. *Algebraic joint zero-diagonalization and blind sources separation*. Signal Processing, IEEE Transactions on, vol. 56, no. 3, pages 980–989, 2008. (Cité en page 19.)
- [Chang 2000] Chunqi Chang, Zhi Ding, Sze Fong Yau et F. H Y Chan. *A matrix-pencil approach to blind separation of colored nonstationary signals*. Signal Processing, IEEE Transactions on, vol. 48, no. 3, pages 900–907, 2000. (Cité en page 9.)
- [Chawla 2009a] MPS Chawla. *Detection of indeterminacies in corrected ECG signals using parameterized multidimensional independent component analysis*. Computational and Mathematical Methods in Medicine, vol. 10, no. 2, pages 85–115, 2009. (Cité en pages 13 et 76.)
- [Chawla 2009b] MPS Chawla. *Multidimensional independent component analysis for statistical estimations of indeterminacies in electrocardiograms*. Journal of Mechanics in Medicine and Biology, vol. 09, no. 03, pages 345–375, 2009. (Cité en pages 13 et 76.)
- [Choi 2000] Seungjin Choi, Andrzej Cichocki et Shun-Ichi Amari. *Flexible independent component analysis*. Journal of VLSI signal processing systems for signal, image and video technology, vol. 26, no. 1-2, pages 25–38, 2000. (Cité en page 76.)
- [Choi 2001] Seungjin Choi, A. Cichocki et A. Belouchrani. *Blind separation of second-order nonstationary and temporally colored sources*. In Statistical Signal Processing, 2001. Proceedings of the 11th IEEE Signal Processing Workshop on, pages 444–447, 2001. (Cité en page 9.)
- [Choi 2002a] S. Choi. *A symmetric-definite pencil approach to source separation*. in : Proceedings of the ITC’2002, vol. CD, 2002. (Cité en page 9.)
- [Choi 2002b] Seungjin Choi et Adel Beloucharni. *Second order nonstationary source separation*. Journal of VLSI Signal Processing, vol. 32, pages 93–104, 2002. (Cité en page 9.)
- [Choi 2006] Heeyoul Choi et Seungjin Choi. *Relative Gradient Learning for Independent Subspace Analysis*. In Neural Networks, 2006. IJCNN ’06. International Joint Conference on, pages 3919–3924, 2006. (Cité en page 14.)
- [Cichocki 1992] A. Cichocki et L. Moszczynski. *New learning algorithm for blind separation of sources*. Electronics Letters, vol. 28, no. 21, pages 1986–1987, 1992. (Cité en page 8.)

- [Cichocki 1996] A. Cichocki et R. Unbehauen. *Robust neural networks with on-line learning for blind identification and blind separation of sources*. Circuits and Systems I : Fundamental Theory and Applications, IEEE Transactions on, vol. 43, no. 11, pages 894–906, 1996. (Cité en page 8.)
- [Cichocki 2007] A. Cichocki, S. Amari, K. Siwek, T. Tanaka, Anh Huy Phan et R. Zdunek. *ICALAB-MATLAB Toolboxes Ver. 3 for signal processing*. Technical report, Laboratory for Advanced Brain Signal Processing, Japan, March 2007. (Cité en page 62.)
- [Comon 1989a] P. Comon. *Separation Of Stochastic Processes*. In Higher-Order Spectral Analysis, 1989. Workshop on, pages 174–179, 1989. (Cité en page 8.)
- [Comon 1989b] P. Comon. *Séparation de mélanges de signaux*. GRETSI, Groupe d’Etudes du Traitement du Signal et des Images, 12° Colloque sur le traitement du signal et des images, FRA, 1989, pages 137–140, 1989. (Cité en page 8.)
- [Comon 1994] P. Comon. *Independent component analysis, A new concept ?* Signal Processing, vol. 36, no. 3, pages 287–314, 1994. (Cité en pages 7, 8, 9, 11, 12 et 23.)
- [Comon 2001] P. Comon. *From source separation to blind equalization, contrast-based approaches*. in ICISP 01, Int. Conf. on Image and Signal Processing, Agadir, Morocco, May 3-5, pages 20–32, 2001. (Cité en pages 9 et 67.)
- [Comon 2007] P. Comon. *Séparation de sources. : Vol. 1 : concepts de base et analyse en composantes indépendantes. Traité IC2 : information, commande, communication. Traitement du signal et de l’image*. Hermes Science Publications, 2007. (Cité en page 7.)
- [Comon 2010] P. Comon et C. Jutten. *Handbook of blind source separation : Independent component analysis and applications*. Independent Component Analysis and Applications Series. Elsevier Science, 2010. (Cité en page 7.)
- [Darmois 1953] George Darmois. *Analyse générale des liaisons stochastiques : etude particulière de l’analyse factorielle linéaire*. Revue de l’Institut international de statistique, pages 2–8, 1953. (Cité en pages 11 et 15.)
- [De Lathauwer 1994] Lieven De Lathauwer, Bart De Moor et Joos Vandewalle. *Blind source separation by higher-order singular value decomposition*. In Proc. EUSIPCO, volume 1, pages 175–178, 1994. (Cité en page 75.)
- [De Lathauwer 1995] Lieven De Lathauwer, D Callaerts, B De Moor et J Vandewalle. *Fetal electrocardiogram extraction by source subspace separation*. In Proc. HOS, volume 95, pages 134–8, 1995. (Cité en pages vii, 75 et 76.)
- [De Lathauwer 1997] Lieven De Lathauwer. *Signal Processing Based on Multilinear Algebra*. PhD thesis, Katholieke Universiteit Leuven, 1997. (Cité en page 10.)
- [De Lathauwer 2001] L. De Lathauwer, B. De Moor et J. Vandewalle. *Independent component analysis and (simultaneous) third-order tensor diagonalization*. Signal Processing, IEEE Transactions on, vol. 49, no. 10, pages 2262–2271, 2001. (Cité en pages 10 et 67.)
- [De Lathauwer 2002] L. De Lathauwer, C. Févotte, B. De Moor et J. Vandewalle. *Jacobi Algorithm for Joint Block Diagonalization in Blind Identification*. In Proc. 23rd Symp. on Information Theory in the Benelux, Louvain-la-Neuve, Belgium, May 2002. (Cité en page 20.)
- [De Lathauwer 2003] L. De Lathauwer, B. De Moor, J. Vandewalle et J.-F. Cardoso. *Independent Component Analysis of Largely Underdetermined Mixtures*. In Proc. 4th Int. Symp. on Independent Component Analysis and Blind Signal Separation (ICA 2003), pages 29–34, Nara, Japan, April 2003. (Cité en page 10.)
- [De Lathauwer 2007] L. De Lathauwer, J. Castaing et J. Cardoso. *Fourth-Order Cumulant-Based Blind Identification of Underdetermined Mixtures*. Signal Processing, IEEE Transactions on, vol. 55, no. 6, pages 2965–2973, 2007. (Cité en page 10.)

- [De Moor 1997] Bart De Moor, P. De Gersem, Bart De Schutter et W. Favoreel. *DAISY : A database for identification of systems*. JOURNAL A, vol. 38, pages 4–5, 1997. (Cit  en pages vii, 76 et 77.)
- [Ding 2006] Shuxue Ding. *A Power Iteration Algorithm for ICA Based on Diagonalizations of Non-Linearized Covariance Matrix*. ICICIC, vol. 2, pages 730–733, 2006. (Cit  en page 60.)
- [Eriksson 2003] Jan Eriksson et Visa Koivunen. *Characteristic-function-based independent component analysis*. Signal Process., vol. 83, no. 10, pages 2195–2208, Octobre 2003. (Cit  en page 15.)
- [Fan 2004] Zhimin Fan, Jie Zhou et Ying Wu. *Motion segmentation based on independent subspace analysis*. 2004. (Cit  en page 13.)
- [Ferreol 1997] A. Ferreol et P. Chevalier. *Higher order blind source separation using the cyclostationarity property of the signals*. In Acoustics, Speech, and Signal Processing, 1997. ICASSP-97., 1997 IEEE International Conference on, volume 5, pages 4061–4064 vol.5, 1997. (Cit  en pages 10 et 67.)
- [Ferreol 2005] A. Ferreol, L. Albera et P. Chevalier. *Fourth-order blind identification of underdetermined mixtures of sources (FOBIUM)*. Signal Processing, IEEE Transactions on, vol. 53, no. 5, pages 1640–1653, 2005. (Cit  en pages 10, 67 et 72.)
- [F votte 2007a] C F votte et F. Theis. *Orthonormal approximate joint block-diagonalization*. Technical report GET/T l com Paris, 2007. (Cit  en pages 14, 15, 20 et 50.)
- [F votte 2007b] C dric F votte et Fabian J. Theis. *Pivot selection strategies in Jacobi joint block-diagonalization*. In Independent Component Analysis and Signal Separation, pages 177–184. Springer, 2007. (Cit  en pages 14 et 20.)
- [Georgiev 2003] P. Georgiev et A. Cichocki. *Robust Independent Component Analysis via Time-Delayed Cumulant Functions*. IEICE Trans, Fundamentals, vol. E86-A, no. 03, pages 573–579, March 2003. (Cit  en page 60.)
- [Ghennoui 2007] Hicham Ghennoui, F. El Mostafa, Nad ge Thirion-Moreau, Abdellah Adib et Eric Moreau. *A nonunitary joint block diagonalization algorithm for blind separation of convolutive mixtures of sources*. Signal Processing Letters, IEEE, vol. 14, no. 11, pages 860–863, 2007. (Cit  en pages 21 et 50.)
- [Ghennoui 2010] Hicham Ghennoui, Nad ge Thirion-Moreau, Eric Moreau et Driss Aboutajdine. *Gradient-based joint block diagonalization algorithms : Application to blind separation of FIR convolutive mixtures*. Signal Processing, vol. 90, no. 6, pages 1836–1849, 2010. (Cit  en pages 21 et 50.)
- [Ghurye 1962] SG Ghurye et Ingram Olkin. *A characterization of the multivariate normal distribution*. The Annals of Mathematical Statistics, vol. 33, no. 2, pages 533–541, 1962. (Cit  en pages 15 et 16.)
- [Gijswijt 2010] Dion Gijswijt. *Matrix algebras and semidefinite programming techniques for codes*. arXiv preprint arXiv :1007.0906, 2010. (Cit  en page 20.)
- [Gorokhov 1997] A. Gorokhov et P. Loubaton. *Subspace-based techniques for blind separation of convolutive mixtures with temporally correlated sources*. Circuits and Systems I : Fundamental Theory and Applications, IEEE Transactions on, vol. 44, no. 9, pages 813–820, 1997. (Cit  en page 7.)
- [Granger 1976] Clive WJ Granger. *Tendency towards normality of linear combinations of random variables*. Metrika, vol. 23, no. 1, pages 237–248, 1976. (Cit  en page 11.)
- [Gretton 2005] Arthur Gretton, Ralf Herbrich, Alexander Smola, Olivier Bousquet et Bernhard Sch lkopf. *Kernel Methods for Measuring Independence*. J. Mach. Learn. Res., vol. 6, pages 2075–2129, D cembre 2005. (Cit  en page 14.)
- [Gruber 2006] Peter Gruber, , Peter Gruber et Fabian J. Theis. *Grassmann Clustering*. In in Proceedings of the European Signal Processing Conference, (EUSIPCO 2006, 2006. (Cit  en page 14.)
- [Gruber 2009] Peter Gruber, Harold W. Gutch et Fabian J. Theis. *Hierarchical Extraction of Independent Subspaces of Unknown Dimensions*. In Proceedings of the 8th International Conference on Independent Component Analysis and Signal Separation, ICA ’09, pages 259–266, Berlin, Heidelberg, 2009. Springer-Verlag. (Cit  en pages 14 et 15.)

- [Gutch 2007] H. Gutch et F. Theis. *Independent Subspace Analysis is unique, given irreducibility*. In Proc. ICA 2007, vol. 4666 of LNCS, pages 49–56, 2007. (Cit  en pages 15, 16, 17, 18, 19, 40 et 53.)
- [Gutch 2011] Harold W Gutch, Jan Krumsiek et Fabian J Theis. *An ISA algorithm with unknown group sizes identifies meaningful clusters in metabolomics data*. In Proc. EUSIPCO, volume 2011, pages 1733–1737, 2011. (Cit  en page 13.)
- [Gutch 2012] Harold W. Gutch et Fabian J. Theis. *Uniqueness of linear factorizations into independent subspaces*. J. Multivar. Anal., vol. 112, pages 48–62, Novembre 2012. (Cit  en pages 15, 16, 17, 18, 40 et 53.)
- [H. Gutch 2010] T. Maehara H. Gutch et F. Theis. *Second Order Subspace Analysis and Simple Decompositions*. In Proc. ICA 2010, vol. 6365 of LNCS, pages 370–377, Springer, St. Malo, France 2010. (Cit  en pages 15, 20 et 39.)
- [Hald 2000] Anders Hald. *The Early History of the Cumulants and the Gram-Charlier Series*. International Statistical Review, vol. 68, no. 2, pages 137–153, 2000. (Cit  en page 23.)
- [Hardy 2006] Michael Hardy. *Combinatorics of Partial Derivatives*. The electronic journal of combinatorics, vol. 13, pages 1–13, 2006. (Cit  en pages 24 et 26.)
- [Haykin 2000] Simon Haykin. *Unsupervised adaptive filtering : Blind source separation*. Adaptive and learning systems for signal processing, communications, and control. Wiley, 2000. (Cit  en page 7.)
- [Hyv rinen 1997] A. Hyv rinen et E. Oja. *A Fast Fixed-Point Algorithm for Independent Component Analysis*. Neural Computation, vol. 9, no. 07, pages 1483–1492, 1997. (Cit  en pages 8, 60 et 67.)
- [Hyvarinen 1999] A. Hyvarinen. *Fast and robust fixed-point algorithms for independent component analysis*. Neural Networks, IEEE Transactions on, vol. 10, no. 3, pages 626–634, 1999. (Cit  en pages 8, 11, 14 et 60.)
- [Hyv rinen 2000] Aapo Hyv rinen et Patrik Hoyer. *Emergence of Phase- and Shift-Invariant Features by Decomposition of Natural Images into Independent Feature Subspaces*. Neural Comput., vol. 12, no. 7, pages 1705–1720, Juillet 2000. (Cit  en pages 12 et 14.)
- [Hyv rinen 2001] Aapo Hyv rinen, Patrik O. Hoyer et Mika O. Inki. *Topographic Independent Component Analysis*. Neural Comput., vol. 13, no. 7, pages 1527–1558, Juillet 2001. (Cit  en pages 12, 13, 14 et 19.)
- [Hyv rinen 2004a] A. Hyv rinen, J. Karhunen et E. Oja. *Independent component analysis*. Adaptive and Learning Systems for Signal Processing, Communications and Control Series. Wiley, 2004. (Cit  en page 7.)
- [Hyv rinen 2004b] Aapo Hyv rinen et Jarmo Hurri. *Blind separation of sources that have spatiotemporal variance dependencies*. Signal Process., vol. 84, no. 2, pages 247–254, F vrier 2004. (Cit  en page 19.)
- [Hyv rinen 2006] A. Hyv rinen et U. K ster. *FastISA : A fast fixed-point algorithm for independent subspace analysis*. pages 371–376, 2006. (Cit  en pages 12 et 14.)
- [H rault 1984] J. H rault et B. Ans. *R seaux neuronaux   synapses modifiables : D codage de messages sensoriels composites par apprentissage non supervis  et permanent*. C. R. de l’Acad mie des Sciences, Paris s rie III, pages 525–528, 1984. (Cit  en page 8.)
- [Ilin 2006] Alexander Ilin. *Independent dynamics subspace analysis*. In Proc. of the 14th European Symposium on Artificial Neural Networks (ESANN 2006), 2006. (Cit  en page 12.)
- [J. H rault 1985] C.Jutten J. H rault et B. Ans. *D tection de grandeurs primitives dans un message composite par une architecture de calcul n oromantique en apprentissage non supervis *. X me colloque GRETSI, Nice, France, 20-24 May, pages 1017–1022, 1985. (Cit  en page 8.)
- [Jinadasa 1986] K. G. Jinadasa et D. S. Tracy. *Higher order moments of random vectors using matrix derivatives*. Stochastic Analysis and Applications, vol. 4, no. 4, pages 399–407, 1986. (Cit  en page 24.)

- [Joho 2002] M. Joho et Kamran Rahbar. *Joint diagonalization of correlation matrices by using Newton methods with application to blind signal separation*. In Sensor Array and Multichannel Signal Processing Workshop Proceedings, 2002, pages 403–407, 2002. (Cité en pages 8 et 19.)
- [Jutten 1988] Christian Jutten. *Independent components analysis versus principal components analysis*. Signal Processing IV, Theories and Applications (EUSIPCO'88) Grenoble, France, pages 643–646, 1988. (Cité en page 8.)
- [Jutten 1991] Christian Jutten et Jeanny Herault. *Blind separation of sources, part I : An adaptive algorithm based on neuromimetic architecture*. Signal Processing, vol. 24, no. 1, pages 1 – 10, 1991. (Cité en page 8.)
- [Jutten 2000] Christian Jutten et Anisse Taleb. *Source Separation : From Dusk Till Dawn*. In Proc. Int. Symp. Independent Component Analysis and Blind Signal Separation, pages 15–26, 2000. (Cité en page 8.)
- [Kagan 1973] A.M. Kagan, I.U.V. Linnik et C.R. Rao. Characterization problems in mathematical statistics. Wiley series in probability and mathematical statistics. Wiley, 1973. (Cité en pages 15, 16 et 46.)
- [Kawanabe 2007a] Motoaki Kawanabe, Masashi Sugiyama, Gilles Blanchard et Klaus-Robert Müller. *A new algorithm of non-Gaussian component analysis with radial kernel functions*. Annals of the Institute of Statistical Mathematics, vol. 59, no. 1, pages 57–75, 2007. (Cité en pages 19 et 67.)
- [Kawanabe 2007b] Motoaki Kawanabe et Fabian J Theis. *Joint low-rank approximation for extracting non-Gaussian subspaces*. Signal Processing, vol. 87, no. 8, pages 1890–1903, 2007. (Cité en pages 19 et 67.)
- [Kendall 1977] M.G. Kendall et A. Stuart. The advanced theory of statistics : Distribution theory, volume 1 of *The Advanced Theory of Statistics*. C. Griffin, 1977. (Cité en page 11.)
- [Khatri 1972] CG Khatri et C Radhakrishna Rao. *Functional equations and characterization of probability laws through linear functions of random variables*. Journal of Multivariate Analysis, vol. 2, no. 2, pages 162–173, 1972. (Cité en pages 15 et 16.)
- [Kim 2003] Heyjin Kim, Seungjin Choi et Sung-Yang Bang. *Membership scoring via independent feature subspace analysis for grouping co-expressed genes*. In Neural Networks, 2003. Proceedings of the International Joint Conference on, volume 3, pages 1690–1695 vol.3, 2003. (Cité en pages 12 et 13.)
- [Kim 2005] Hyejin Kim. *S. : Independent subspaces of gene expression data*. In In : Proc. IASTED Int ?1 Conf. Artificial Intelligence and Applications, 2005. (Cité en page 13.)
- [Kim 2006] Jong Kyoung Kim et Seungjin Choi. *Tree-Dependent components of gene expression data for clustering*. In Proceedings of the 16th international conference on Artificial Neural Networks - Volume Part II, ICANN'06, pages 837–846, Berlin, Heidelberg, 2006. Springer-Verlag. (Cité en page 13.)
- [Kirshner 2008] Sergey Kirshner et Barnabás Póczos. *ICA and ISA using Schweizer-Wolff measure of dependence*. In Proceedings of the 25th international conference on Machine learning, ICML '08, pages 464–471, New York, NY, USA, 2008. ACM. (Cité en page 15.)
- [Kohonen 1996] Teuvo Kohonen. *Emergence of invariant-feature detectors in the adaptive-subspace self-organizing map*. Biological Cybernetics, vol. 75, no. 4, pages 281–291, 1996. (Cité en page 14.)
- [Koldovský 2006] Z. Koldovský, P. Tichavský et E. Oja. *Efficient Variant of Algorithm FastICA for Independent Component Analysis Attaining the Cramér-Rao Lower Bound*. IEEE Trans. on Neural Networks, vol. 17, no. 05, pages 1265–1277, Sept 2006. (Cité en page 8.)
- [Kollo 1991] T. Kollo. Matrix derivative in multivariate statistics. Tartu University Press. Tartu (in Russian), 1991. (Cité en page 86.)

- [Kollo 2005] T. Kollo et D. von Rosen. *Advanced multivariate statistics with matrices*. Springer, 2005. (Cit  en pages 11 et 23.)
- [Kollo 2008] T. Kollo. *Multivariate skewness and kurtosis measures with an application in ICA*. *Journal of Multivariate Analysis*, vol. 99, pages 2328–2338, 2008. (Cit  en pages 9, 23, 24, 41, 42, 68 et 69.)
- [Lahat 2009] Dana Lahat, J.F. Cardoso et Hagit Messer. *Optimal Performance of Second-Order Multidimensional ICA*. *Intelligent Data Analysis-IDA*, no. 50–57, 2009. (Cit  en pages 14 et 39.)
- [Lahat 2012a] D. Lahat, J. Cardoso et H. Messer. *Second-Order Multidimensional ICA : Performance Analysis*. *Signal Processing, IEEE Transactions on*, vol. 60, no. 9, pages 4598–4610, 2012. (Cit  en pages 14, 15 et 39.)
- [Lahat 2012b] Dana Lahat, Jean-Fran ois Cardoso et Hagit Messer. *Joint Block Diagonalization Algorithms for Optimal Separation of Multidimensional Components*. In Fabian Theis, Andrzej Cichocki, Arie Yeredor et Michael Zibulevsky,  diteurs, *Latent Variable Analysis and Signal Separation*, volume 7191 of *Lecture Notes in Computer Science*, pages 155–162. Springer Berlin Heidelberg, 2012. (Cit  en pages 14 et 21.)
- [Le 2011] Quoc V. Le, Will Y. Zou, Serena Y. Yeung et Andrew Y. Ng. *Learning hierarchical invariant spatio-temporal features for action recognition with independent subspace analysis*. In *CVPR’11*, pages 3361–3368, 2011. (Cit  en page 13.)
- [Lee 1998] Te W. Lee, Mark Girolami, Anthony J. Bell et Terrence J. Sejnowski. *A unifying Information-theoretic framework for Independent Component Analysis*, 1998. (Cit  en page 11.)
- [Lee 1999] Te-Won Lee, Mark Girolami et Terrence J. Sejnowski. *Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources*. *Neural Comput.*, vol. 11, no. 2, pages 417–441, F vrier 1999. (Cit  en page 8.)
- [Leonov 1959] VP Leonov et AN Shiryaev. *On a method of calculation of semi-invariants*. *Theory of Probability & its applications*, vol. 4, no. 3, pages 319–329, 1959. (Cit  en page 23.)
- [Li 2001a] S.Z. Li, Xiao Guang Lv et HongJiang Zhang. *View-based clustering of object appearances based on independent subspace analysis*. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 2, pages 295–300 vol.2, 2001. (Cit  en page 13.)
- [Li 2001b] S.Z. Li, Xiao Guang Lv et HongJiang Zhang. *View-subspace analysis of multi-view face patterns*. In *Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, 2001. Proceedings. IEEE ICCV Workshop on*, pages 125–132, 2001. (Cit  en page 13.)
- [Li 2010] Rui Li, Hongwei Li et Fasong Wang. *Dependent Component Analysis : Concepts and Main Algorithms*. *Journal of Computers*, vol. 5, no. 4, 2010. (Cit  en pages 13 et 76.)
- [Llinares 2010] R. Llinares, J. Igual, J. Miro-Borras et A. Camacho. *Atrial activity estimation using periodic component analysis*. In *Neural Networks (IJCNN), The 2010 International Joint Conference on*, pages 1–7, 2010. (Cit  en page 19.)
- [Loperfido 2014] Nicola Loperfido. *A note on the fourth cumulant of a finite mixture distribution*. *Journal of Multivariate Analysis*, vol. 123, no. 0, pages 386–394, 1 2014. (Cit  en pages 24 et 68.)
- [Lukacs 1955] Eugene Lukacs. *Applications of Faa Di Bruno’s Formula in Mathematical Statistics*. *The American Mathematical Monthly*, vol. 62, no. 5, pages 340–348, May 1955. (Cit  en pages 24 et 26.)
- [Ma 2010] Sai Ma, Xi-Lin Li, N.M. Correa, T. Adali et V.D. Calhoun. *Independent subspace analysis with prior information for fMRI data*. In *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pages 1922–1925, 2010. (Cit  en page 13.)
- [Mac Cullagh 1987] P. Mac Cullagh. *Tensor methods in statistics*. *Monographs on Statistics and Applied Probability*. Chapman and Hall, 1987. (Cit  en pages 11, 23, 26, 37 et 41.)
- [MacRae 1974] E.C. MacRae. *Matrix derivatives with an application to an adaptive linear decision problem*. *Ann. Statist*, no. 337–346, 1974. (Cit  en pages 45 et 86.)

- [Maehara 2010] T. Maehara et K. Murota. *A numerical algorithm for block-diagonal decomposition of matrix *-algebras with general irreducible components*. Japan Journal of Industrial and Applied Mathematics, vol. 27, pages 263–293, 2010. (Cit  en pages 20 et 50.)
- [Maehara 2011] T. Maehara et K. Murota. *Algorithm for error-controlled simultaneous block-diagonalization of matrices*. SIAM Journal on Matrix Analysis and Applications, vol. 32, pages 605–620, 2011. (Cit  en pages 15, 20, 39 et 50.)
- [Magnus 1979] Jan R. Magnus et H. Neudecker. *The Commutation Matrix : Some Properties and Applications*. Annals of Statistics, vol. 7, no. 2, pages 381–394, March 1979. (Cit  en pages 24, 27, 35, 36 et 85.)
- [Marossero 2003] Dorothee E Marossero, Deniz Erdogmus, Neil Euliano, Jose C Principe, KE Hildet *al*. *Independent components analysis for fetal electrocardiogram extraction : a case for the data efficient mermaid algorithm*. In Neural Networks for Signal Processing, 2003. NNSP’03. 2003 IEEE 13th Workshop on, pages 399–408. IEEE, 2003. (Cit  en page 76.)
- [Meinecke 2002] F. Meinecke, A. Ziehe, M. Kawanabe et K. Muller. *A resampling approach to estimate the stability of one-dimensional or multidimensional independent components*. Biomedical Engineering, IEEE Transactions on, vol. 49, no. 12, pages 1514–1525, 2002. (Cit  en pages 13 et 76.)
- [Mendel 1991] J. M. Mendel. *Tutorial on higher-order statistics (spectra) in signal processing and system theory : theoretical results and some applications*. Proceedings of the IEEE In Proceedings of the IEEE, vol. 79, no. 3, pages 278–305, March 1991. (Cit  en page 43.)
- [Moreau 2001] E. Moreau. *A generalization of joint-diagonalization criteria for source separation*. Signal Processing, IEEE Transactions on, vol. 49, no. 3, pages 530–541, mar 2001. (Cit  en pages 10, 19, 52 et 60.)
- [M ri 1993] Tam s F M ri, Vijay K Rohatgi et GJ Sz kely. *On multivariate skewness and kurtosis*. Theory of Probability & Its Applications, vol. 38, no. 3, pages 547–551, 1993. (Cit  en page 68.)
- [Murota 2010] K. Murota, Y. Kanno, M. Kojima et S. Kojima. *A numerical algorithm for block-diagonal decomposition of matrix *-algebras with application to semidefinite programming*. Japan Journal of Industrial and Applied Mathematics, vol. 27, pages 125–160, 2010. (Cit  en pages 20, 39 et 50.)
- [Neudecker 1969] H. Neudecker. *Some Theorems on Matrix Differentiation with Special Reference to Kronecker Matrix Products*. Journal of the American Statistical Association, vol. 64, no. 327, pages 953–963, 1969. (Cit  en pages 23 et 36.)
- [Nion 2011] Dimitri Nion. *A tensor framework for nonunitary joint block diagonalization*. Signal Processing, IEEE Transactions on, vol. 59, no. 10, pages 4585–4594, 2011. (Cit  en pages 21 et 50.)
- [Nishimori 2006a] Yasunori Nishimori, Shotaro Akaho et Mark D. Plumbley. *Riemannian Optimization Method on Generalized Flag Manifolds for Complex and Subspace ICA*. AIP Conference Proceedings, vol. 872, no. 1, pages 89–96, 2006. (Cit  en page 14.)
- [Nishimori 2006b] Yasunori Nishimori, Shotaro Akaho et Mark D. Plumbley. *Riemannian optimization method on the flag manifold for independent subspace analysis*. In Proceedings of the 6th international conference on Independent Component Analysis and Blind Signal Separation, ICA’06, pages 295–302, Berlin, Heidelberg, 2006. Springer-Verlag. (Cit  en page 14.)
- [Nishimori 2007] Y. Nishimori, S. Akaho, S. Abdallah et M.D. Plumbley. *Flag Manifolds for Subspace ICA Problems*. In Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on, volume 4, pages IV–1417–IV–1420, 2007. (Cit  en page 14.)
- [Nishimori 2008] Yasunori Nishimori, Shotaro Akaho et Mark D. Plumbley. *Natural Conjugate Gradient on Complex Flag Manifolds for Complex Independent Subspace Analysis*. In Proceedings of the 18th international conference on Artificial Neural Networks, Part I, ICANN ’08, pages 165–174, Berlin, Heidelberg, 2008. Springer-Verlag. (Cit  en page 14.)

- [Nordhausen 2011] Klaus Nordhausen et Hannu Oja. *Independent subspace analysis using three scatter matrices*. Austrian Journal of Statistics. v40 i1-2, pages 93–101, 2011. (Cité en pages 15 et 39.)
- [Nordhausen 2012] Klaus Nordhausen, Harold W. Gutch, Hannu Oja et Fabian J. Theis. *Joint diagonalization of several scatter matrices for ICA*. In Proceedings of the 10th international conference on Latent Variable Analysis and Signal Separation, LVA/ICA'12, pages 172–179, Berlin, Heidelberg, 2012. Springer-Verlag. (Cité en page 15.)
- [Ould-Baba 2011] H. Ould-Baba, J. Antoni et V. Robin. *Two methods of blind source separation exploiting only the higher order statistics*. The International Conference Surveillance 6, October 25-26, 2011 at the University of Technology of Compiègne, France, pages 1–12, 2011. (Cité en pages 10 et 71.)
- [Ould-Baba 2013] H. Ould-Baba, V. Robin et J. Antoni. *Concise formulae for the cumulant matrices of a random vector*. submitted, 2013. (Cité en page 51.)
- [Peña 2010] Daniel Peña, Francisco J Prieto et Júlia Viladomat. *Eigenvectors of a kurtosis matrix as interesting directions to reveal cluster structure*. Journal of Multivariate Analysis, vol. 101, no. 9, pages 1995–2007, 2010. (Cité en page 68.)
- [Petrochilos 2006] Nicolas Petrochilos et Pierre Comon. *A zero-cumulant random variable and its applications*. Signal processing, vol. 86, no. 11, pages 3334–3338, 2006. (Cité en page 52.)
- [Pham 1997] D.-T. Pham et P. Garat. *Blind separation of mixture of independent sources through a quasi-maximum likelihood approach*. Signal Processing, IEEE Transactions on, vol. 45, no. 7, pages 1712–1725, 1997. (Cité en page 14.)
- [Pham 2000] Dinh Tuan Pham. *Joint Approximate Diagonalization of Positive Definite Hermitian Matrices*. SIAM J. Matrix Anal. Appl., vol. 22, no. 4, pages 1136–1152, Juillet 2000. (Cité en pages 8 et 19.)
- [Pham 2001] Dinh-Tuan Pham et J.-F. Cardoso. *Blind separation of instantaneous mixtures of nonstationary sources*. Signal Processing, IEEE Transactions on, vol. 49, no. 9, pages 1837–1848, 2001. (Cité en page 14.)
- [Pierre Comon 1991] Christian Jutten Pierre Comon et Jeanny Herault. *Blind separation of sources, part II : Problems statement*. Signal Processing, vol. 24, no. 1, pages 11 – 20, 1991. (Cité en page 8.)
- [Póczos 2005a] Barnabás Póczos et András Lörincz. *Independent subspace analysis using geodesic spanning trees*. In Proceedings of the 22nd international conference on Machine learning, ICML '05, pages 673–680, New York, NY, USA, 2005. ACM. (Cité en page 14.)
- [Póczos 2005b] Barnabás Póczos et András Lörincz. *Independent subspace analysis using k-nearest neighborhood distances*. In Proceedings of the 15th international conference on Artificial neural networks : formal models and their applications - Volume Part II, ICANN'05, pages 163–168, Berlin, Heidelberg, 2005. Springer-Verlag. (Cité en pages 14 et 64.)
- [Qi 2007] Liqun Qi, Wenyu Sun et Yiju Wang. *Numerical multilinear algebra and its applications*. Frontiers of Mathematics in China, vol. 2, no. 4, pages 501–526, 2007. (Cité en page 28.)
- [Roberts 2001] S. Roberts et R. Everson. *Independent component analysis : Principles and practice*. Cambridge University Press, 2001. (Cité en page 7.)
- [S. Cruces 2001] S. Amari S. Cruces A. Cichocki. *The Minimum Entropy and Cumulant Based Contrast Functions for Blind Source Extraction*. in Lecture Notes in Computer Science, Springer-Verlag, IWANN'2001, vol. II, pages 786–793, 2001. (Cité en page 60.)
- [Samarov 2002] Alexander Samarov et Alexandre Tsybakov. *Nonparametric Independent Component Analysis*, 2002. (Cité en page 14.)
- [Sameni 2008] Reza Sameni, Christian Jutten et Mohammad B Shamsollahi. *Multichannel electrocardiogram decomposition using periodic component analysis*. Biomedical Engineering, IEEE Transactions on, vol. 55, no. 8, pages 1935–1940, 2008. (Cité en page 19.)

- [Santos 2005] C.S. Santos, Jr. Kogler S.E. et E. Del Hernandez. *Using independent subspace analysis for selecting filters used in texture processing*. In Image Processing, 2005. ICIP 2005. IEEE International Conference on, volume 1, pages I-465–8, 2005. (Cité en page 13.)
- [Saul 2001] Lawrence K Saul et Jont B Allen. *Periodic component analysis : an eigenvalue method for representing periodic structure in speech*. Advances in Neural Information Processing Systems, pages 807–813, 2001. (Cité en page 19.)
- [Schott 2003] J. R. Schott. *Kronecker product permutation matrices and their application to moment matrices of the normal distribution*. Journal of Multivariate Analysis, vol. 87, pages 177–190, 2003. (Cité en pages 24, 35 et 36.)
- [Sharma 2006] Alok Sharma et Kuldeep K Paliwal. *Subspace independent component analysis using vector kurtosis*. Pattern Recognition, vol. 39, no. 11, pages 2227–2232, 2006. (Cité en pages 12, 13, 14, 15 et 76.)
- [Skitovič 1953] V.P. Skitovič. *On a property of the normal distribution*. (Russian) Doklady Akad. Nauk SSSR (N.S.), vol. 89, pages 217–219, 1953. (Cité en pages 11 et 15.)
- [Skitovich 1954] Viktor Pavlovich Skitovich. *Linear forms of independent random variables and the normal distribution law*. Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya, vol. 18, no. 2, pages 185–200, 1954. (Cité en page 15.)
- [Smaragdis 1998] Paris Smaragdis. *Blind separation of convolved mixtures in the frequency domain*. Neurocomputing, vol. 22, no. 1, pages 21–34, 1998. (Cité en page 7.)
- [Soon 1990] V.C. Soon, L. Tong, Y. Huang et R. Liu. *An extended fourth order blind identification algorithm in spatially correlated noise*. In Acoustics, Speech, and Signal Processing, 1990. ICASSP-90., 1990 International Conference on, pages 1365–1368 vol.3, 1990. (Cité en page 9.)
- [Sorensen 2008] Mikael Sorensen, Sylvie Icart, Pierre Comon, Luc Deneire et al. *Gradient based approximate joint diagonalization by orthogonal transforms*. In 16th European Signal Processing Conference (Eusipco'08), 2008. (Cité en pages 8 et 19.)
- [Sorouchyari 1991] E. Sorouchyari. *Blind separation of sources, part III : Stability analysis*. Signal Processing, vol. 24, no. 1, pages 21 – 29, 1991. (Cité en page 8.)
- [Stögbauer 2004] H. Stögbauer, A. Kraskov, S.A. Astakhov et P. Grassberger. *Least-dependent-component analysis based on mutual information*. Phys Rev E Stat Nonlin Soft Matter Phys, vol. 70, no. 6 Pt 2, page 066123, 2004. (Cité en page 14.)
- [Szabó 2007] Zoltán Szabó, Barnabás Póczos et András Lórinicz. *Undercomplete Blind Subspace Deconvolution*. J. Mach. Learn. Res., vol. 8, pages 1063–1095, Décembre 2007. (Cité en page 14.)
- [Theis 2004a] F. J. Theis. *Uniqueness of complex and multidimensional independent component analysis*. Signal Process., vol. 84, no. 5, pages 951–956, Mai 2004. (Cité en pages 12, 14, 15 et 16.)
- [Theis 2004b] Fabian J. Theis. *A New Concept for Separability Problems in Blind Source Separation*, 2004. (Cité en page 15.)
- [Theis 2004c] Fabian J Theis, Anke Meyer-Baese et Elmar W Lang. *Second-order blind source separation based on multi-dimensional autocovariances*. In Independent Component Analysis and Blind Signal Separation, pages 726–733. Springer, 2004. (Cité en page 9.)
- [Theis 2005a] F. Theis. *Blind signal separation into groups of dependent signals using joint block diagonalization*. pages 5878 – 5881 Vol. 6, may 2005. (Cité en pages vii, 12, 13, 14, 15, 39, 64 et 76.)
- [Theis 2005b] Fabian J. Theis. *Multidimensional independent component analysis using characteristic functions*. In In Proc. of EUSIPCO, 2005. (Cité en pages 12, 14, 15 et 39.)
- [Theis 2006] Fabian J Theis et Motoaki Kawanabe. *Uniqueness of non-gaussian subspace analysis*. In Independent Component Analysis and Blind Signal Separation, pages 917–925. Springer, 2006. (Cité en pages 19 et 67.)

- [Theis 2007a] F. Theis. *Towards a general independent subspace analysis*. In Proc. NIPS 2006, pages 1361–1368, 2007. (Cit  en pages vii, 15, 16, 17, 18, 20, 39, 40, 41, 50, 51, 53, 63, 64 et 71.)
- [Theis 2007b] FabianJ. Theis et M. Kawanabe. *Colored Subspace Analysis*. In MikeE. Davies, ChristopherJ. James, SamerA. Abdallah et MarkD Plumbley,  diteurs, Independent Component Analysis and Signal Separation, volume 4666 of *Lecture Notes in Computer Science*, pages 121–128. Springer Berlin Heidelberg, 2007. (Cit  en page 19.)
- [Theis 2010] F.J. Theis. *Colored Subspace Analysis : Dimension Reduction Based on a Signal’s Autocorrelation Structure*. Circuits and Systems I : Regular Papers, IEEE Transactions on, vol. 57, no. 7, pages 1463–1474, 2010. (Cit  en page 19.)
- [Tian 2005] Yongge Tian. *Some rank equalities and inequalities for Kronecker products of matrices*. Linear and Multilinear Algebra, vol. 53, no. 6, pages 445–454, 2005. (Cit  en page 54.)
- [Tichavsky 2012] P Tichavsky et Z Koldovsky. *Algorithms for nonorthogonal approximate joint block-diagonalization*. In Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European, pages 2094–2098. IEEE, 2012. (Cit  en page 21.)
- [Tkacenko 2001] Andre Tkacenko et PP Vaidyanathan. *Generalized kurtosis and applications in blind equalization of MIMO channels*. In Signals, Systems and Computers, 2001. Conference Record of the Thirty-Fifth Asilomar Conference on, volume 1, pages 742–746. IEEE, 2001. (Cit  en page 68.)
- [Tom  2006] Ana Maria Tom . *The generalized eigendecomposition approach to the blind source separation problem*. Digit. Signal Process., vol. 16, no. 3, pages 288–302, Mai 2006. (Cit  en page 9.)
- [Tong 1990] L. Tong, V.C. Soon, Y. Huang et R. Liu. *AMUSE : a new blind identification algorithm*. pages 1784–1787 vol.3, 1990. (Cit  en page 9.)
- [Tong 1991] Lang Tong, R. Liu, V.C. Soon et Yih-Fang Huang. *Indeterminacy and identifiability of blind identification*. Circuits and Systems, IEEE Transactions on, vol. 38, no. 5, pages 499–509, 1991. (Cit  en pages 7 et 8.)
- [Tong 1992] Lang Tong, Yujiro Inouye et R-W Liu. *A finite-step global convergence algorithm for the parameter estimation of multichannel MA processes*. Signal Processing, IEEE Transactions on, vol. 40, no. 10, pages 2547–2558, 1992. (Cit  en page 72.)
- [Tracy 1993] D. S. Tracy et S. A. Sultan. *Higher order moments of multivariate normal distribution using matrix derivatives*. Stochastic Analysis and Applications, vol. 11, no. 3, pages 337–348, 1993. (Cit  en pages 24 et 36.)
- [Vigneron 2003] Vincent Vigneron, Anisoara Paraschiv-Ionescu, Annabelle Azancot, O Sibony et C Jutten. *Fetal electrocardiogram extraction based on non-stationary ICA and wavelet denoising*. In Signal Processing and Its Applications, 2003. Proceedings. Seventh International Symposium on, volume 2, pages 69–72. IEEE, 2003. (Cit  en page 76.)
- [Vollgraf 2001] Roland Vollgraf et Klaus Obermayer. *Multi-dimensional ICA to separate correlated sources*. In Neural Information Processing Systems. Cambridge, MA. MIT Press, vol. 14, pages 993–1000, 2001. (Cit  en page 14.)
- [von B nau 2009] Paul von B nau, Frank C Meinecke et Klaus-Robert M ller. *Stationary subspace analysis*. In Independent Component Analysis and Signal Separation, pages 1–8. Springer, 2009. (Cit  en page 19.)
- [Ye 2008] Yalan Ye, Zhi-Lin Zhang, Jiazhi Zeng et Lei Peng. *A fast and adaptive ICA algorithm with its application to fetal electrocardiogram extraction*. Applied Mathematics and Computation, vol. 205, no. 2, pages 799–806, 2008. (Cit  en page 76.)
- [Yeredor 2000] Arie Yeredor. *Blind source separation via the second characteristic function*. Signal Processing, vol. 80, no. 5, pages 897–902, 2000. (Cit  en page 15.)

- [Yeredor 2002] A. Yeredor. *Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation*. Signal Processing, IEEE Transactions on, vol. 50, no. 7, pages 1545–1553, 2002. (Cité en pages 8 et 19.)
- [Yeredor 2004] Arie Yeredor, Andreas Ziehe et Klaus-Robert Müller. *Approximate Joint Diagonalization Using a Natural Gradient Approach*. In CarlosG. Puntonet et Alberto Prieto, éditeurs, Independent Component Analysis and Blind Signal Separation, volume 3195 of *Lecture Notes in Computer Science*, pages 89–96. Springer Berlin Heidelberg, 2004. (Cité en pages 8 et 19.)
- [Yonghong 2012] XIE Yonghong. *Application research of fetal ECG signal extraction based on MICA technology*. Computer Engineering and Applications, vol. 48, no. 14, pages 235–240, 2012. (Cité en pages 13 et 76.)
- [Zaidi 2005] Abdelhamid Zaidi. *Positive definite combination of symmetric matrices*. Signal Processing, IEEE Transactions on, vol. 53, no. 11, pages 4412–4416, 2005. (Cité en page 72.)
- [Zarzoso 2009] Vicente Zarzoso. *Extraction of ECG Characteristics Using Source Separation Techniques : Exploiting Statistical Independence and Beyond*. In Amine Nait-Ali, éditeur, Advanced Biosignal Processing, pages 15–47. Springer Berlin Heidelberg, 2009. (Cité en page 76.)
- [Ziehe 1998] Andreas Ziehe et Klaus-Robert Müller. *TDSEP - an efficient algorithm for blind separation using time structure*. 1998. (Cité en page 9.)
- [Ziehe 2004] Andreas Ziehe, Pavel Laskov, Guido Nolte et Klaus-Robert Müller. *A fast algorithm for joint diagonalization with non-orthogonal transformations and its application to blind source separation*. The Journal of Machine Learning Research, vol. 5, pages 777–800, 2004. (Cité en pages 8, 19 et 53.)
- [Zoltan Szabo 2012] Andras Lorincz Zoltan Szabo Barnabas Poczos. *Separation Theorem for Independent Subspace Analysis and its Consequences*. Pattern Recognition, 2012. (Cité en pages 11 et 60.)